BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 4 :
C12N 15/00, 7/00, 1/20
C12N 1/18, 5/00, C07K 15/04
A61K 39/155

A1

(11) International Publication Number: WO 87/ 04185

(43) International Publication Date: 16 July 1987 (16.07.87)

(21) International Application Number: PCT/US86/02756

(22) International Filing Date: 23 December 1986 (23.12.86)

(31) Priority Application Number: 818,740

(32) Priority Date: 14 January 1986 (14.01.86)

(33) Priority Country: US

(60) Parent Application or Grant
(63) Related by Continuation
US 818,740 (CIP)
Filed on 14 January 1986 (14.01.86)

(71) Applicant *(for all designated States except US)*: UNIVERSITY OF NORTH CAROLINA [US/US]; Chapel Hill, NC 27514 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)* : WERTZ, Gail, W. [US/US]; 608 Laurel Hill Road, Chapel Hill, NC 27514 (US). COLLINS, Peter, L. [US/US]; 4703 Saul Road, Kensington, MD 20895 (US).

(74) Agent: WEBER, Kenneth, A.; Patent Law Department, The Upjohn Company, Kalamazoo, MI 49001 (US).

(81) Designated States: AT (European patent), AU, BE (European patent), CH (European patent), DE (European patent), DK, FI, FR (European patent), GB (European patent), IT (European patent), JP, KR, LU (European patent), NL (European patent), NO, SE (European patent), US.

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: VACCINES FOR HUMAN RESPIRATORY VIRUS

(57) Abstract

Compositions of DNA and protein that are useful for preparing vaccines against human respiratory syncytial virus (HRSV). The proteins include the native structural viral proteins and immunogenic fragments thereof. The DNA compositions include structural genes coding for these proteins and expression and replication plasmids containing the structural genes. Host cells transformed with the above DNA compositions are also disclosed herein. Lastly vaccines comprised of the native structural viral proteins and their immunogenic derivatives are disclosed as well as methods for protecting humans by inoculation with said vaccines.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**FR**	France	**ML**	Mali
AU	Australia	**GA**	Gabon	**MR**	Mauritania
BB	Barbados	**GB**	United Kingdom	**MW**	Malawi
BE	Belgium	**HU**	Hungary	**NL**	Netherlands
BG	Bulgaria	**IT**	Italy	**NO**	Norway
BJ	Benin	**JP**	Japan	**RO**	Romania
BR	Brazil	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SN**	Senegal
CH	Switzerland	**LI**	Liechtenstein	**SU**	Soviet Union
CM	Cameroon	**LK**	Sri Lanka	**TD**	Chad
DE	Germany, Federal Republic of	**LU**	Luxembourg	**TG**	Togo
DK	Denmark	**MC**	Monaco	**US**	United States of America
FI	Finland	**MG**	Madagascar		

VACCINES FOR HUMAN RESPIRATORY VIRUS

BACKGROUND OF THE INVENTION

Field of the Invention.

This invention discloses compositions of DNA and protein that are useful for preparing vaccines against human respiratory syncytial virus [HRSV]. The proteins include the native structural viral proteins and immunogenic fragments thereof. The DNA compositions include structural genes coding for these proteins and expression and replication plasmids containing the structural genes. Host cells transformed with the above DNA compositions are also disclosed herein. Lastly vaccines comprised of the native structural viral proteins and their immunogenic derivatives are disclosed as well as methods for protecting humans by inoculation with said vaccines. This invention was made with Government support under At-12464 awarded by the National Institute of Health. The Government has certain rights in the invention.

Background.

HRSV was first discovered in 1956 and is worldwide in distribution. It is an important cause of upper and lower respiratory tract disease causing illness in infants and young children. In infants this severe illness often requires hospitalization. About 30 percent of hospitalized young children with acute respiratory disease have respiratory syncytial virus infection. In older children and adults the disease is milder. Infections with respiratory syncytial virus are referable to all segments of the respiratory tract, are usually associated with fever, cough, runny nose, and fatigue, and are diagnosed clinically as bronchitis, bronchiolitis, pneumonia, croup, or viral infection. In older children and adults the virus is generally limited to replication in the upper respiratory tract. Infants may be more severely involved when the virus extends into the lungs. Lung damage can be permanent.

Primary infection with respiratory syncytial virus occurs early in life, usually before 4 years of age. Among children, illness caused by this virus tends to occur at least once each year in rather sharply defined outbreaks of several months duration. Epidemics are sharply circumscribed, generally for 3 to 5 months. In family studies, children in early school years frequently introduce the virus into the home, infecting younger members of the family more

severely than other family members. The clinical consequence of infection is most severe on first experience and becomes milder in older individuals who are immunologically experienced.

The effects of respiratory syncytial virus can range from inapparent infection to severe pneumonia and death. Inflammation of the respiratory track is responsible for most symptoms. Complete recovery in most cases occurs in one to three weeks with the production of antibody which appears to persist throughout life. In the United States about 30 percent of 1-year old infants and 95 percent of 5-year old children have circulating respiratory syncytial virus antibody. Reinfections in older infants, children, and adults with antibody are mostly mild upper respiratory illnesses in the form of colds.

With exception of the present invention, there are no effective vaccines to combat HRSV.

Description of the Prior Art

Although low yields of virus in cell culture have hindered HRSV research, the virus has been well studied. HRSV is a paramyxovirus containing a single negative strand of RNA which is transcribed into 10 predominantly monocistronic messengers. The messengers have been isolated and translated in vitro. The products have been characterized by gel electrophoresis, peptide mapping and immuno-precipitation as being similar to structural proteins isolated from virions. The structural proteins include a major nucleocapsid protein (N; MW ca. 42,000), a nucleocapsid phosphoprotein (P; MW ca. 34,000), a large nucleocapsid protein (L; MW ca. 200,000), an envelope matrix protein (M; MW ca. 26,000), a matrix glycoprotein (ca. 22,000) and two envelope glycoproteins, the fusion glycoprotein (F; MW ca. 68,000 to 70,000) and a second, methioninepoor glycoprotein (G; MW ca. 84,000 to 90,000). In addition, a virally encoded protein of about 9,500 daltons and other small proteins are known to be present in infected cells. Collins, P.L., et al., Identification of a tenth mRNA of RSV and assignment of polypeptides to the 10 viral genes-,J. of Virol. 49:572 -578 (1984) and references cited therein. Although the structural proteins of HRSV have been isolated, their amino acid sequences are not known.

Multiple attempts have been made to obtain an effective vaccine against HRSV. Friedewald et al., Journal of the American Medical

Association, Vol. 204, 20 May 1968, pp. 690-694 describe the propagation of respiratory syncytial virus in bovine embryonic kidney tissue culture. Virus grown at 34°C or 28°C did not decrease in infectivity or virulence. HRSV grown at 26°C, while associated with a decrease in infectivity for adults, could not be considered for use in prevention of infection in adults since the virus had limited infectivity and was poorly immunogenic.

Kim et al., Pediatrics, Vol. 48, November 1971, pp. 745-755, disclose that inactivated respiratory syncytial virus vaccine prepared from virus grown at 26°C stimulated the development of high levels of serum antibody in infants and children from 6 months to 13 years in age but did not prevent infection.

McIntosh et al., Pediatric Research, Vol. 8, 1974, pp. 689-696, discuss two experimental live respiratory syncytial virus vaccines, one prepared from virus grown at 26°C. and the other, prepared from a temperature sensitive mutant which grew well at 32°C and not at all at 37°C. or higher. The first vaccine was unsatisfactory as it did not protect against infection when the interval between vaccination and challenge was greater than 4 months. The second vaccine was also unsatisfactory in that it apparently lost its temperature sensitivity in some vaccinees.

Craighead, Journal of Infectious Diseases, Vol. 131, June 1975, pp. 749-753, discusses tests conducted in 1966 wherein several groups of investigators tested in infants and young children a formaldehyde-treated, alum-precipitated virus grown in tissue culture. Upon subsequent exposure to wild virus the vaccine recipients exhibited an accentuated pattern of respiratory tract disease. Craighead concludes that immunization with formaldehyde treated virus enhanced the severity of the disease.

Wright et al., Journal of Pediatrics, Vol. 88, June 1976, pp. 931-936, describe the evaluation in infants of a temperature sensitive live attenuated respiratory syncytial vaccine. While this vaccine when administered at a dosage level sufficiently high to infect all seronegative infants caused mild upper respiratory illness, lowering the dose did not achieve an acceptable level of infectivity. The virus was also genetically unstable as there was evidence of loss of temperature sensitivity in one vaccinee. There

was no evidence for potentiation of natural illness with this vaccine and reinfection occurred among vaccinees.

US patent Nos. 4,122,167 and 4,145,252 describe a method for attenuating virions by serial passage through human diploid lung fibroblasts and US patent No. 4,517,304 discloses a method for producing immunogenically active HRSV proteins upon the cell membranes of susceptible cells grown in culture. These cells are then injected into a host to elicit an immune response.

None of the above references disclose the methods or compositions disclosed in this invention. The above references attempt to create a vaccine by injection of virions comprised of both protein and nucleic acid or by injection of undefined compositions of virus proteins attached to the cell membranes of host cells. None of the above work has resulted in an effective vaccine. Raeburn, P., The Houdini Virus, Science 85, Vol 6:52-57 (Dec. 1985). Disclosed herein are compositions of pure viral protein and methods for producing commercially practical amounts of that protein. The viral proteins are useful for producing vaccines, antibodies for diagnostics, and the clones carrying the HRSV-like cDNA can also be used for diagnostic purposes. Moreover, vaccines produced from the proteins can be tailored to contain any proportion of the structural proteins that will best afford immuno-protection. From a safety perspective this invention avoids the exposure of young children to intact HRSV virions either inactivated or attenuated and to viral nucleic acid. By avoiding the injection of a complete virion, the vaccines disclosed herein need not be treated with a fixative such as formaldehyde which has been shown to result in the development of ineffective antibodies and in the subsequent increased susceptibility of the host/patient when exposed to virulent HRSV.

The following references by the inventors of this invention are offered to complete the relevant HRSV literature, but are not prior art references under 35 U.S.C. 102(b): (1) Collins, P.L. et al., Nucleotide Sequence of the gene encoding the fusion (F) glycoprotein of human respiratory syncytial virus, Proc. Natl. Acad. Sci., USA, 81:7683-7687 (December 1984) disclosing the gene sequence for the F glycoprotein; (2) Collins, P.L. et al., The 1A Protein Gene of Human Respiratory Syncytial Virus: Nucleotide Sequence of the mRNA and a Related Polycistronic Transcript, Virology, 141:283-291 (1985)

disclosing the gene sequence for the 1A protein; (3) Collins, P.L. et al., The Envelope-Associated 22K Protein of Human Respiratory Syncytial Virus: Nucleotide Sequence of the mRNA and a Related Polytranscript, J. of Virol., 54(No.1):65-71 (Apr. 1985) disclosing the gene sequence for the 22K protein; (4) Wertz, G.W. et al., Nucleotide sequence of the G protein gene of human respiratory syncytial virus reveals an unusual type of viral membrane protein, Proc. Natl. Acad. Sci., USA, 82:4075-4079 (June 1985) disclosing the gene sequence for the G glycoprotein; and (5) Collins, P.L. et al., Correct Sequence for the Major Nucleocapsid Protein mRNA of Respiratory Syncytial Virus, Virology, 146:69-77 (1985) disclosing the gene sequence for the N protein.

In 1986, it was demonstrated that the vaccinia virus expression system was useful for expressing the G and F glycoproteins of HRSV. Ball, L.A., et al, Expression of the Major Glycoprotein G of Human Respiratory Syncytial Virus from Recombinant Vaccinia Virus Vectors, P.N.A.S. USA 83:246-250 (1986) and Olmsted, R.A., Expression of the F Glycoprotein of Respiratory Syncytial Virus by a Recombinant Vaccinia Virus: Comparison of the Individual Contributions of the F and G Glycoproteins to Host Immunity, P.N.A.S. USA 83:7462-7466 (1986). These two glycoproteins were also demonstrated to induce immunoprotection in mammals against a live HRSV virus challenge. Stott, E.J., et al., Human Respiratory Syncytial Virus Glycoprotein G Expressed from Recombinant Vaccinia Virus Vector Protects Mice Against Live-virus Challenge, Journal of Virology 67:607-613 (1986); Elango N., et al., Resistance and Human Respiratory Syncytial Virus (RSV) Infection Induced by Immunization of Cotton Rats with a Recombinant Vaccinia Virus Expressing the RSV G Glycoprotein; and, Olmsted, R.A. (supra) P.N.A.S. USA 83:246-250 (1986). The methodology and results of the above references are all incoporated by reference herein.

SUMMARY OF THE INVENTION

This invention relates to the development of a vaccine for protecting humans from HRSV. The vaccine is comprised of a composition of viral structural proteins and a suitable carrier. The proteins are produced through expression of recombinant DNA by suitable cell hosts. Plasmids, cDNA sequences, transformed cell hosts, and vaccines are disclosed herein.

Specifically, this invention discloses a DNA sequence coding for human respiratory syncytial virus structural proteins selected from the group consisting of: F protein, G protein, 22 K protein, 9.5 K protein; major capsid protein N and immunogenic fragments thereof. Most preferred are the G and F glycoproteins and immunogenic fragments thereof.

In addition there are disclosed compositions of DNA sequences coding for the above HRSV structural proteins or immunogenic fragments wherein the sequence is recombined into a plasmid capable of independent replication in a suitable host, of incorporation into the host genome or of inducing expression of the DNA sequences coding for viral proteins or fragments in a suitable host. Suitable hosts include bacteria, yeast and eukaryote cell cultures.

This invention also discloses compositions of essentially pure protein selected from the group of HRSV structural proteins consisting of: F protein, G protein, 22 K protein, 9.5 K protein; major capsid protein N and immunogenic fragments thereof.

Vaccines and methods of using the vaccines are disclosed herein in which the vaccine is comprised of a polypeptide selected from the group of HRSV structural proteins consisting of: F protein, G protein, 22 K protein, 9.5 K protein; major capsid protein N and immunogenic fragments thereof. Most preferred are the F protein, G protein and immunogenic fragments thereof.

Detailed Description

This invention involves a series of molecular genetic manipulations that can be achieved in a variety of known ways. The following descriptions will detail the various methods available to express the HRSV proteins and are followed by specific examples of preferred methods.

In summary the manipulations can be described as the obtaining of a cDNA of HRSV proteins, the cloning and replication of the cDNA in E. coli and the expression of the desired cDNA in a suitable host.

The specific sequence and base numbering positions for the disclosed proteins of HRSV strain A_2 are illustrated in Charts 12-16. Charts 12-16 contain the nucleic acid sequences for HRSV structural proteins F protein, G protein, 22 K protein, 9.5 K protein, and major capsid protein N.

It is anticipated that protein from the A_2 strain will induce cross-protection against other strains of HRSV; however, it is possible that maximum protection will involve immunization with a mixture of proteins from various strains.

A. General Methods

Generally, the nomenclature and general laboratory procedures required in this application can be found in Maniatis, T. et al., Molecular Cloning A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1982. The manual is hereinafter referred to as Maniatis.

All E. coli strains are grown on Luria broth (LB) with glucose, Difco's Antibiotic Medium #2 and M9 medium supplemented with glucose and acid-hydrolyzed casein amino acids. Strains with resistance to antibiotics were maintained at the drug concentrations described in Maniatis. Transformations were performed according to the method described by Rowekamp, W. and Firtel, R.A., Dev. Biol., 79:409-418 (1980).

All enzymes were used according to the manufacturer's instructions. Transformants were analyzed by colony hybridization as described in Grunstein, M. and Wallis, J., Methods in Enzymology, 68:379-388.

After hybridization, the probes are removed and saved, and the filters are washed in 0.1% SDS, 0.2x SSC for a total of 3 hours with 5 changes of 400 ml each. Filters are thoroughly air dried, mounted, and autoradiographed using Kodak X-OMAT AR film and Dupont Cronex Lightnening Plus intensifying screens for 16 hours at -70° C.

For sequencing of plasmids, purified plasmid DNA is prepared according to the methods described in Maniatis. End-labeled DNA fragments are prepared and analyzed by the chemical sequencing methods of Maxam and Gilbert with modifications described by Collins, P.L. and Wertz, G.W., J. Virol. 54:65-71 (1985).

Nucleotide sizes are given in either kilobases (kb) or basepairs (bp). These are estimates derived from agarose gel electrophoresis.

B. HRSV cDNA

The first step in obtaining expression of HRSV proteins is to obtain the DNA sequence coding for the protein from cDNA clones. This sequence is then cloned into an expression plasmid which is

capable of directing transcription of the gene and allowing efficient translation of the transcript.

The library method for obtaining cDNA encoding HRSV proteins has been described generally in Maniatis, T., Fritsh, E.F., and Sambrook, J. (1982). Molecular Cloning, A Laboratory Manual. Cold Spring Harbor Laboratory, New York.and specifically in Collins, P.L. and Wertz, G.W., cDNA cloning and transcriptional mapping of nine polyadenylated RNAs encoded by the genome of HRSV, Proc. Natl.-Acad. USA 80:3208-3212 (1983).

Clones are prepared by inserting the cDNA into PstI cleaved pBR322 to which homopolymer tracts of dGTP have been enzymatically added to the 3'ends at the cleavage site. Homopolymer tracts of dCTP are enzymatically added to the 3' termini of the cDNA molecules according to the methods described by Maniatis. Ideally, 10-30 residues of dCTP or dGTP should be added to maximize cloning efficiency. The cDNA and plasmid are annealed together and transformed into E. coli. The clones containing full length HRSV cDNA are detected by probes of labeled viral cDNA or oligonucleotides complementary to portions of the sequences illustrated in Charts 12-16, followed by restriction enzyme analysis and DNA sequencing.

Oligonucleotides are chemically synthesized according to the solid phase phosphoramidite triester method first described by Beaucage S.L. and Caruthers, M.H. Tetrahedron Letts. 22(20):1859-1862 (1981) using an automated synthesizer, as described in Needham-VanDevanter, D.R., et al., Nucleic Acids Res., 12:6159-6168 (1984). Purification of oligonucleotides is by either native acrylamide gel electrophoresis or by anion-exchange HPLC as described in Pearson, J.D. and Regnier, F.E., J. Chrom., 255:137-149 (1983).

The sequence of the synthetic oligonucleotides can be verified using the chemical degradation method of Maxam, A.M. and Gilbert, W., Grossman, L. and Moldave, D., eds., Academic Press, New York, Methods in Enzymology, 65:499-560 (1980).

C. Expression in E. coli.

To obtain high level expression of a cloned gene, e.g., the HRSV protein cDNA, in a prokaryotic system, it is essential to construct expression vectors which contain, at the minimum, a strong promoter to direct mRNA transcription, a ribosome binding site for translational initiation, and a transcription terminator. Examples of

regulatory regions suitable for this purpose are the promoter and operator region of the E. coli tryptophan biosynthetic pathway as described by Yanofsky, C., Kelley, R.L. and Horn, V., J. Bacteriol., 158:1018-1024 (1984) and the leftward promoter of phage lambda (P_L) as described by Herskowitz, I. and Hagen, D., Ann. Rev. Genet., 14:399-445 (1980).

The HRSV-like proteins produced in E. coli will not fold properly due to the presence of cysteine residues and to the lack of suitable post-translational modifications. During purification from E. coli, the expressed proteins must first be denatured and then renatured. This can be accomplished by solubilizing the E. coli produced proteins in guanidine HCl and reducing all the cysteine residues with β-mercapto-ethanol. The protein is then renatured either by slow dialysis or by gel filtration. US Patent No. 4,511,503.

Detection of HRSV-like proteins is achieved by methods known in the art such as radioimmunoassays, or Western blotting techniques or immunoprecipitation. Purification from E. coli can be achieved following procedures described in US Patent No. 4,511,503.

D. Expression of HRSV-like proteins in Yeast.

Expression of heterologous proteins in yeast is well known and described. Methods in Yeast Genetics, Sherman, F., et al., Cold Spring Harbor Laboratory, (1982) is a well recognized work describing the various methods used to produce HRSV-like proteins in yeast.

For high level expression of a gene in yeast, it is essential to connect the gene to a strong promoter system as in the prokaryote and to also provide efficient transcription termination/polyadenylation sequences from a yeast gene. Examples of useful promoters include GAL1,10 (Johnston M., and Davis, R.W., Mol. and Cell. Biol., 4:-1440-48, 1984), ADH2 (Russell, D., et al., J. Biol. Chem. 258:-2674-2682, 1983), PHO5 (EMBOJ. 6:675-680, 1982), and MFα1. A multicopy plasmid with a selective marker such as Lue-2, URA-3, Trp-1, and His-3 is also desirable. The MFα1 promoter is preferred. The MFα1 promoter, in a host of the α mating-type is constitutive, but is off in diploids or cells with the a mating-type. It can, however, be regulated by raising or lowering temperature in hosts which have a ts mutation at one of the SIR loci. The effect of such a mutation at 35°C on an α type cell is to

turn on the normally silent gene coding for the a mating-type. The expression of the silent a mating-type gene, in turn, turns off the MFα1 promoter. Lowering the temperature of growth to 27°C reverses the whole process, i.e., turns the a mating-type off and turns the MFα1 on (Herskowitz, I. & Oshima, Y. (1982) in The molecular biology of the yeast saccharomyces, (eds. Strathern, J.N., Jones, E.W., & Broach, J.R., Cold Spring Harbor Lab., Cold Spring Harbor, NY, pp 181-209).

The polyadenylation sequences are provided by the 3'-end sequences of any of the highly expressed genes, like ADH1, MFα1, or TPI (Alber, T. and Kawasaki, G., J. of Mol. & Appl. Genet. 1:419-434, 1982).

A number of yeast expression plasmids like YEp6, YEp13, YEp24 can be used as vectors. A gene of interest such as HRSV-like protein cDNA can be fused to any of the promoters mentioned above, and then ligated to the plasmids for expression in various yeast hosts. The above-
mentioned plasmids have been fully described in the literature (Botstein, et al., Gene, 8:17-24, 1979; Broach, et al., Gene, 8:121-133, 1979).

Two procedures are used in transforming yeast cells. In one case, yeast cells are first converted into protoplasts using zymolyase, lyticase or glusulase, followed by addition of DNA and polyethylene glycol (PEG). The PEG-treated protoplasts are then regenerated in a 3% agar medium under selective conditions. Details of this procedure are given in the papers by J.D. Beggs, Nature (London), 275:104-109 (1978); and Hinnen, A., et al., Proc. Natl. Acad. Sci. USA, 75:1929-1933 (1978). The second procedure does not involve removal of the cell wall. Instead the cells are treated with lithium-chloride or acetate and PEG and put on selective plates (Ito, H., et al., J. Bact., 153:163-168, 1983).

HRSV-like proteins can be isolated from yeast by lysing the cells and applying standard protein isolation techniques to the lysates. The proteins can be detected by using Western blot techniques or radioimmunoassays.

E. Expression in Cell Cultures.

The HRSV cDNA can be ligated to various expression vectors for use in transforming host cell cultures. The vectors all contain gene

sequences to initiate transcription and translation of the HRSV-like proteins that are compatible with the host cell to be transformed.

In addition, the vectors preferably contain a marker to provide a phenotypic trait for selection of transformed host cells such as dihydrofolate reductase or metallothionein. Additionally a replicating vector might contain a replicon.

Illustrative of cell cultures useful for the production of HRSV-like proteins are cells of insect or mammalian origin. Mammalian cell systems often will be in the form of monolayers of cells although mammalian cell suspensions may also be used. Illustrative examples of mammalian cell lines include VERO and HeLa cells, Chinese hamster ovary (CHO) cell lines, WI38, BHK, COS-7 or MDCK cell lines.

As indicated above, the vector, e.g., a plasmid, which is used to transform the host cell preferably contains gene sequences to initiate the transcription and translation of the HRSV-like proteins gene sequence. These sequences are referred to as expression control sequences. When the host cell is of mammalian or insect origin illustrative useful expression control sequences are obtained from the SV-40 promoter (Science, 222, 524-527, 1983), the CMV I.E. promoter (Proc. Natl. Acad. Sci. 81:659-663, 1984), the metallothionein promoter (Nature, 296, 39-42, 1982) or the baculovirus polyhedrin promoter (insect cells) (Virol., 131, 561-565, 1983). The plasmid or replicating or integrating DNA material containing the expression control sequences is cleaved using restriction enzymes and adjusted in size as necessary or desirable and ligated with cDNA coding for HRSV-like proteins by means well known in the art.

As with yeast when higher animal host cells are employed, polyadenylation or transcription terminator sequences from known mammalian genes need to be incorporated into the vector. An example of a terminator sequence is the polyadenylation sequence from the bovine growth hormone gene.

The HRSV glycoprotein F may be designed to be secreted from cells into the surrounding media. This is accomplished by causing the early termination of the glycoprotein prior to its anchor region. L. Lasky, et al., Biotechnology, 2:527-532 (1984). The anchor is a hydrophobic region at the carboxy terminal end of the glycoprotein which causes the retention of the glycoprotein in the cell membrane.

Early termination may be accomplished by inserting a universal translational terminator oligonucleotide into an appropriate site in the gene's DNA. These oligonucleotides are commercially available. For the F gene, a preferred site for insertion is the NsiI restriction enzyme site which is approximately 1.5 kb from the 5' end of the gene.

Additionally gene sequences to control replication in the host cell may be incorporated into the vector such as those found in bovine papillomavirus type-vectors. Saveria-Campo, M., "Bovine papillomavirus DNA: a eukaryotic cloning vector" in DNA Cloning Vol II a practical approach. Ed. D.M. Glover, IRL Press, Arlington, Virginia pages 213-238 (1985).

The preferred expression vector useful for expressing HRSV-like proteins in Chinese hamster ovary (CHO) cell is a shuttle vector pSVCOW7 which replicates in both CHO and E. coli cells utilizing ampicillin resistance and dihydrofolate reductase genes as markers in E. coli and CHO cells respectively. Plasmid pSVCOW7 also provides the polyadenylation sequence from bovine growth hormone which is necessary for expression in CHO cells. Plasmid pSVCOW7 is cleaved and a viral promoter and the HRSV-like protein cDNAs inserted.

The preferred expression vector useful in forming recombinant baculovirus for expressing HRSV-like proteins in insect cells is pAc373. Smith et al., Mol. Cell. Biol. 3:2156-2165 (1983). The plasmid replicates in E. coli cells utilizing ampicillin resistance, and provides the eukaryotic promoter and polyadenylation signal from the baculovirus polyhedrin gene for expression of HRSV genes. Plasmid pAc373 is cleaved and a HRSV cDNA is inserted adjacent to the promoter. This new plasmid is cotransfected with baculovirus (Autograpa californica nuclear polyhedrosis virus) DNA into insect cells by calcium phosphate precipitation. Recombinant baculovirus in which the pAc373 polyhedrin gene containing a HRSV cDNA has replaced the resident viral polyhedrin gene by homologous recombination is detected by dot blot hybridization (Summers, M., and G, Smith, A manual of methods for baculovirus vectors and insect cell culture procedures, Texas A & M University, College Station, Texas, pp. 29-30 (1986)) using ^{32}p-labeled HRSV cDNA as a probe. Insect cells infected with recombinant baculovirus may also be differentiated by their inclusion-negative morphology since the insertion of the HRSV

cDNA into the polyhedrin gene prevents the synthesis of this inclusion-forming protein. Isolation of HRSV proteins from infected insect cells is accomplished as described for CHO cells.

The preferred expression vector used in conjunction with bovine papilloma virus (BPV) for expressing HRSV-like proteins is pTFW9 U.S. Serial No. 935,490, which is incorporated by reference herein. The plasmid replicates in E. coli utilizing ampicillin resistance, and provides the mouse metallothionein promoter and SV40 polyadenylation signal for expression of HRSV genes. Plasmid pTFW9 is cleaved and a HRSV cDNA is inserted adjacent to the promoter. This new plasmid is then cleaved to allow insertion of BPV. The recombinant plasmid is transfected into animal cells by calcium phosphate precipitation and foci of transformed cells are selected. HRSV protein expressed in these transformed cells is isolated as described for CHO cells.

The host cells are competent or rendered competent for transfection by various means. There are several well-known methods of introducing DNA into animal cells. These include: calcium phosphate precipitation, fusion of the recipient cells with bacterial protoplasts containing the DNA, treatment of the recipient cells with liposomes containing the DNA, and microinjection of the DNA directly into the cells.

The transfected cells are cultured by means well known in the art. Biochemical Methods in Cell Culture and Virology, Kuchler, R. J., Dowden, Hutchinson and Ross, Inc., (1977). and the expressed HRSV-like proteins analogs are isolated from cell suspensions created by disruption of the host cell system by well known mechanical or enzymatic means. HRSV-like proteins which are designed to be secreted from the cells are isolated from the media without disruption of the cells.

Isolation of the HRSV proteins is accomplished by lysing the CHO cells with detergents. For HRSV glycoproteins it is helpful to first apply the cytoplasmic fraction to a lentil lectin column which will specifically bind glycoproteins. The eluted glycoproteins are then applied to an affinity column containing anti-HRSV antibody. Non-glycoproteins of HRSV can be directly applied to the affinity column.

F. Definitions.

The phrase "cell culture" refers to the containment of growing cells derived from either a multicellular plant or animal which allows for the cells to remain viable outside the original plant or animal.

The term "downstream" identifies sequences proceeding farther in the direction of expression; for example, the coding region is downstream from the initiation codon.

The term "microorganism" includes both single cellular prokaryote and eukaryote organisms such as bacteria, actinomycetes and yeast.

The term "operon" is a complete unit of gene expression and regulation, including structural genes, regulator genes and control elements in DNA recognized by regulator gene product.

The term "plasmid" refers to an autonomous self-replicating extrachromosomal circular DNA and includes both the expression and nonexpression types. Where a recombinant microorganism or cell culture is described as hosting an expression plasmid the phrase "expression plasmid" includes both extrachromosomal circular DNA and DNA that has been incorporated into the host chromosome(s). Where a plasmid is being maintained by a host cell, the plasmid is either being stably replicated by the cells during mitosis as an autonomous structure or as an incorporated portion of the host's genome.

The term "promoter" is a region of DNA involved in binding the RNA polymerase to initiate transcription.

The phrase "immunogenic fragments" includes derivatives of the structural proteins of HRSV having sufficient antigenic capacity to produce effective immunologic protection in patient exposed to virulent HRSV. The phrase "HRSV-like proteins" is meant to encompass these fragments. For example, HRSV proteins are made up of amino acid residues, not all of which are exposed to the aqueous environment and capable of eliciting a strong immunogenic response. If carefully selected, modification or deletion to these regions would not affect antigenicity. While no longer being native HRSV proteins, the proteins are now immunogenic fragments if deletions are involved and HRSV-like proteins if either deletions or modifications to the primary sequence were involved.

The phrase "DNA sequence" refers to a single or double stranded

DNA molecule comprised of nucleotide bases, adenosine, thymidine, cytosine and guanosine.

The phrase "essentially pure (HRSV) protein" refers to compositions of viral protein that contain no virus synthesized protein. Although the essentially pure proteins may be contaminated with low levels of host cell constituents, the protein is devoid of contaminating structural and non-structural viral protein produced by replicating HRSV.

The phrase "suitable host" refers to a cell culture or microorganism that is compatible with a recombinant plasmid and will permit the plasmid to replicate, to be incorporated into its genome or to be expressed.

The term "upstream" identifies sequences proceeding in the opposite direction from expression; for example, the bacterial promoter is upstream from the transcription unit, the initiation codon is upstream from the coding region.

Conventions used to represent plasmids and fragments in Charts 1- 6, are meant to be synonymous with conventional circular representations of plasmids and their fragments. Unlike the circular figures, the single line figures on the charts represent both circular and linear double-stranded DNA with initiation or transcription occurring from left to right (5' to 3'). Asterisks (*) represent the bridging of nucleotides to complete the circular form of the plasmids. Fragments do not have asterisk marks because they are linear pieces of double-stranded DNA. Endonuclease restriction sites are indicated above the line. Gene markers are indicated below the line. Bars appearing below the diagrams representing the plasmid or fragments are used to indicate the number of basepairs between two points on the DNA. The relative spacing between markers do not indicate actual distances but are only meant to indicate their relative positions on the illustrated DNA sequence.

EXAMPLES

Example 1. The Cloning of HRSV Glycoproteins F and G.

A. Virus and Cells.

The A_2 strain of RS virus, (available from the American Type Culture Collection, Bethesda, MD), is propagated in monolayer cultures of HEp-2 cells in Eagle minimum essential medium supplemented with 5% heat-inactivated fetal calf serum. Viral infectivity

is measured by cytopathic effect on monolayer cultures of HEp-2 cells.

B. Preparation of Radiolabeled RS Virus Intracellular RNAs.

Monolayer cultures of HEp-2 cells are infected with RS virus at a multiplicity of infection of 1 PFU per cell. After 2 hours of adsorption at 37°C, fresh Eagle minimal essential medium supplemented with 5% heat-inactivated fetal calf serum is added. At 14 hours postinfection (p.i.), the cells are treated with 5 μg of actinomycin D per ml. The cells are then exposed to [^{3}H]uridine at 20 μCi/ml in the presence of drug from 16 to 20 h p.i.

C. Preparation of Purified HRSV mRNA's.

At 20 hours postinfection, cells are suspended in HBS solution (10 mM N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid (HEPES), pH 7.6, 10 mM NaCl, 1 mM $MgCl_2$) and broken by Dounce homogenization. Nuclei are removed by centrifugation at 2,000 x g. The supernatant is made approximately 4.5 M with respect to CsCl and 1.5% in N-lauryl sarcosine and is layered over 2 ml of 5.7 M CsCl solution containing HBS, 0.1 M EDTA, and 2% N-lauryl sarcosine. After 12 to 24 h of centrifugation in a Beckman SW40 rotor at 25,000 rpm and 22°C, the clear RNA pellet is resuspended in sterile water, brought to 0.2 M NaCl-0.2% sodium dodecyl sulfate (SDS), and ethanol precipitated. After a second precipitation with ethanol, mRNA's are isolated by binding to oligodeoxythymidylate [oligo(dT)]-cellulose in 0.01 M Trishydrochloride, pH 7.5, containing 0.02% SDS and 0.5 M NaCl, and eluting in the above minus the NaCl. Eluted mRNA's are precipitated with ethanol after addition of rabbit liver tRNA carrier and NaCl to 0.2 M.

D. cDNA Synthesis.

The synthesis of cDNA follows conditions designed to maximize cDNA length. Land, H. et al. Nuc. Acids Res. 9:2251-2266 (1981). Twenty-five micrograms of poly(A)$^+$ RNA from RS virus-infected cells is transcribed into cDNA by using 40 μg of oligo(dT) as primer and 140 units of avian myeloblastosis virus reverse transcriptase (Life Sciences, St. Petersburg, FL) in a 500-μl reaction mixture containing: Tris.HCl (50 mM, pH 8.3); $MgCl_2$ (10 mM); dithiothreitol (30 mM); KCl (120 mM); sodium pyrophosphate (4 mM); dTTP, dATP, and dGTP (1 mM each); [^{3}H]dCTP (ICN Radiochemicals, 0.8 mCi, 0.4 Ci/mmol; 1 Ci=3.7 x 10^{10} becquerels); and $(dT)_{12-18}$ (80 μg/ml); and mRNA (50

μg/ml). The mixture is incubated for 1 hr at 43°C and the reaction is terminated by phenol-chloroform extraction and ethanol precipitation.

The nucleic acids are resuspended in water and incubated for 2 h at 37°C in the presence of 0.3 M NaOH (final volume, 300 μl). The mixture is neutralized by the additions of 25 μl of 2.5 M Tris-hydrochloride (pH 7.6) and 30 μl of 2 M HCl and is immediately passed through Sephadex G-200 with a column buffer of 1 mM Tris-hydrochloride (pH 7.6). The cDNAs contained in the leading edge of the void volume are collected. Homopolymeric dCMP tails are added in a 550-μl reaction mixture containing 325 units of terminal transferase (P-L Biochemicals). The reaction mixture is incubated at 15°C. Aliquots are withdrawn after 2.5 and 5 min and adjusted to 10 mM EDTA, and the cDNAs are purified by extraction with phenol-chloroform, followed by three rounds of ethanol precipitation.

Synthesis of the second cDNA strand is performed in a 600-μl reaction mixture under the conditions described above for reverse transcription of mRNA, except the actinomycin D is omitted, the oligodeoxythymidylate is replaced by 30 μg of oligodeoxyguanylate$_{12-18}$ (P-L Biochemicals) per ml, and the reaction contains 0.75 mCi of [α^{-32}P]dCTP (specific activity, 3,000 Ci/mmol; Amersham Corp.). After incubation for 1 h at 43°C, the reaction mixture is passed directly through Sepharose 6B, and the cDNAs in the void volume are recovered. To obtain maximum completion of second-strand synthesis, the cDNAs are placed in a 400-μl reaction mixture containing 10 mM Tris-hydrochloride (pH 7.6), 8 mM magnesium acetate, 70 mM KCl, 10 mM dithioerythritol, 0.5 mM each deoxynucleotide, and 12 units of DNA polymerase I (Klenow fragment) (P-L Biochemicals). After incubation for 2 h at 15°C, the reaction is terminated by the addition of EDTA to 10 mM. The products are purified by extraction with phenol-chloroform and passage through Sepharose 6B. Homopolymer dCMP tails are added in a 600-μl reaction mixture under the conditions described above, except incubations take place at 30°C for 2.5 and 5 min. The reactions are terminated by the addition of EDTA and by extraction with phenol-chloroform, and the products are collected by ethanol precipitation.

E. Tailing and Annealing of the cDNA to Vector DNA.

Vector DNA (prepared by digesting pBR322 to completion with PstI and adding homopolymer tracts of dGTP residues) is commercially available from New England Nuclear. The vector DNA can also be made according to the methods described above. The procedure for annealing cDNA with vector DNA is also described by Maniatis. Briefly, tailed cDNA is mixed with vector in a 1:1 molar ratio in a 50 μl reaction containing 10 mM Tris pH 7.4; 0.4 M NaCl; 1 mM EDTA. Final DNA concentrations varied between 20-60 μg/ml. Annealing is accomplished by either; 1) following a defined regimen of incubations consisting of 65°/10'; 42°/60'; 37°/2 hours, and then room temperature for 2 hours, or 2) incubation at 65°/10' shutting off the water bath and allowing it to slowly equilibrate to room temperature overnight.

The cDNA containing vectors are introduced into E. coli using transformation procedures already described. The bacteria are screened in situ using the hybridization procedures also described earlier.

Radioactively labeled [$^{32}_{P}$] hybridization probes are prepared by either of the following methods. The probes may be prepared by reverse transcription of infected cell mRNA which has been prehybridized with uninfected cell mRNA to remove the cellular RNA, or by reverse transcription viral RNA isolated from purified nucleocapsids. Collins, P.L. and Wertz, G., Proc. Natl. Acad. Sci. USA, 80:3208-3212 (1983). Identification of specific cDNAs are achieved by hybrid selection, cell-free translation and immunoprecipitation as described in Collins, P.L. et al., J. Virol. 49(2):572-578 (1984).

The preferred method for colony hybridizations utilizes the sequences disclosed herein to construct the pentadecamers described below as probes. For use as a hybridization probe one μg of 15-mer is phosphorylated in a 50 μl reaction volume consisting of 70 mM Tris-base (pH 7.6), 100 mM KCl; 10 mM $MgCl_2$, 5 mM dithiothreitol, 50 μCi γ^{32}P dATP (P. L. Biochemicals), and 1 U T_4 polynucleotide kinase (New England Biolabs). Incubation is at 37° for 60 minutes. In this fashion, the 15-mer can be labeled to a specific activity of 1 x 10^8 cpm per μg.

F. Plasmids pGPF (chart 1) and pGPG.

Clones exhibiting complementary sequences to the probes complementary to the 5' region of the F and G glycoproteins are selected

for secondary screening using PstI restriction analysis of the clones to determine if the digestion products are consistent with the PstI restriction map which can be obtained from the sequences given in Figures 1-4.

As final proof, a mini-preparation of DNA is isolated from the clone and is sequenced by dideoxy chain termination. Minipreps of plasmid DNA are prepared as described in the General Methods section. Dideoxy sequencing is carried out as described in the General Methods section using the synthetic pentadecamers, described below, as primers in a 20:1 molar excess over template.

G. Synthesizing oligonucleotides complementary to HRSV-like proteins.

Oligonucleotides are chemically synthesized according to the solid phase phosphoramidite triester method first described by Beaucage S.L. and Caruthers, M.H. Tetrahedron Letts. 22(20):1859-1862 (1981) using an automated synthesizer, as described in Needham-VanDevanter, D.R., et al., Nucleic Acids Res., 12:6159-6168 (1984). Purification of oligonucleotides was by either native acrylamide gel electrophoresis or by anion-exchange HPLC as described in Pearson, J.D. and Regnier, F.E., J. Chrom., 255:137-149 (1983).

The sequence of the synthetic oligonucleotides can be verified using the chemical degradation method of Maxam, A.M. and Gilbert, W., Grossman, L. and Moldave, D., eds., Academic Press, New York, Methods in Enzymology, 65:499-560 (1980). Alternatively, the sequence can be confirmed after the assembly of the oligonucleotide fragments into the double-stranded DNA sequence using the method of Maxam and Gilbert, supra, or the chain termination method for sequencing double-stranded templates of Wallace, R.B., et al., Gene, 16:21-26 (1981).

The oligonucleotides from the 3'-end of the mRNA can be used to specifically prime the reverse transcription reaction for making the first strand of the cDNA. The oligonucleotides from the 5'-ends can be used to probe for full length cDNA specific for that gene. The following 15-mer oligonucleotides are useful for the above purposes although alternative sequences could be used.

Ggp

5'-end	ATGTCCAAAAACAAG
3'-end	ACACCACGCCAGTAG

Fgp

5'-end ATGGAGTTGCTAATC

3'-end GCATTTAGTAACTAA

1A

5'-end ATGGAAAATACATCC

3'-end CGAGTCAACACATAG

Nuc

5'-end ATGGCTCTTAGCAAA

3'-end GATGTAGAGCTTTGA

22K

5'-end ATGTCACGAAGGAAT

3'-end AATGATACTACCTGA

Oligonucleotides for Use in λ Exonuclease Step

Fgp

CAAATAACAATGGAG

Ggp

CAAACATGTCCAAAA

Example 2. Expression of Glycoproteins F (gpF) and G (gpG) of HRSV in CHO Cells.

The same procedures and enzymes will be used for both glycoproteins unless otherwise noted.

In order to obtain maximum expression of the F glycoprotein, the G-C nucleotides which are used to insert the cDNA into the plasmid pBR322 must be removed from the 5' end (relative to the original mRNA) of the cDNA. In order to conveniently insert the gpF cDNA into the preferred expression vector for CHO cells, pSVCOW7 (described below), it is necessary to supply a BamHI site upstream from the protein coding sequence. To accomplish this the cDNA of F or G glycoprotein is inserted into pUC12 (PL Pharmacia Labs, Piscataway, N.J.).

A. Construction of pGPF2 - Chart 2.

The cDNA of the glycoproteins is flanked by PstI sites (Chart 1), however there are also internal PstI sites. Therefore, the plasmid pGPF is partially digested with PstI and fragment 2 (1.9 kb; gpG cDNA is 0.9 kb) is isolated from a gel. Fragment 2 is ligated to the plasmid pUC12 (Bethesda Res. Labs,, Rockville, MD) which had been digested with PstI. A plasmid with the 5' end of the gpF gene adjacent to the XbaI site in pUC12 is selected and designated pGPF2

(4.6 kb). This orientation is verified by cleavage with AccI which generates a fragment of approximately 200 bp (for gpG, orientation is verified by digestion with HincII, generating a fragment of approximately 400 bp).

B. Construction of pGPF3 - Chart 3.

To remove the G-C nucleotides from the 5' end of the cDNA, pGPF2 is opened with XbaI and the ends are treated with bacterial alkaline phosphatase to yield fragment 4. Fragment 4 is then digested with SalI which cuts off a small piece between the XbaI and PstI sites and treated with Klenow enzyme to make the ends flush. After treatment with Klenow enzyme, fragment 2 is digested with Lambda exonuclease which requires a 5' phosphate and leaves a 3' overhang. Because of the removal of the 5' phosphate on the end upstream from the gpF, the exonuclease will digest downstream toward the gpF sequence. The exonuclease is allowed sufficient time to remove nucleotides beyond the G/C tail region into the leader sequence. A synthetic sequence containing the first 15 bases of the leader sequence is hybridized to fragment 4 and the missing bases filled in with Klenow enzyme and the ends ligated with T4 ligase to yield pGPF3 (4.6 kb) which is transformed into E. coli and its sequence verified.

To remove the G-C nucleotides from the 3' end of the cDNA, pGPF3 is opened with HindIII and treated with the exonuclease Bal 31 for a time sufficient to digest through the G-C nucleotides. The ends are made blunt with Klenow enzyme and the cDNA clone is freed from the vector DNA by digestion with BamHI. The cDNA fragment is isolated from a gel and ligated to plasmid pUC12 which has been digested with BamHI and HincII (HincII is compatible with blunt ends) to yield pGPF4. The plasmid is transformed into E. coli and an appropriate clone which was sufficiently digested with Bal31 is identified by sequencing. Alternatively, the G-C nucleotides may be removed by digesting with a restriction enzyme which has a unique site upstream from the G-C nucleotides. For gpF such an enzyme whould be HaeIII and for gpG FokI. These ends would be made flush and the DNA treated as described above for generating pGPF4. Since these enzymes cleave upstream from their gene's normal translation termination signal, a universal translation termination oligonucleotide (New England Biolabs) would be ligated into an appropriate restriction enzyme site.

C. Construction of pSVCOW7 - Chart 4.

The starting plasmid pSV2dhfr (available from the American Type Culture Collection or prepared according to the procedure of S. Subramani,et al., "Expression of the Mouse Dihydrofolate Reductase Complementary Deoxyribonucleic Acid in Simian Virus 40", Molecular and Cellular Biology 2:854-864 (Sept. 1981) is digested with BamHI and EcoRI to yield the fragment (5) (5.0 kb) containing the ampicillin resistance gene, the SV40 origin, and the dhfr gene. The second portion of pSVCOW7 is obtained from plasmid pλGH2R2 which is digested with the same restriction endonucleases used to cleave pSV2dhfr to obtain fragment 5 (2.1 kb) containing the 3' end of genomic bovine growth hormone gene, i.e., BGH gDNA. Plasmid pλGH2R2 is publicly available from an E. coli HB101 host, deposited with the Northern Regional Research Laboratories in Peoria, Illinois (NRRL B-15154). Fragments (5 and 6) are ligated to yield pSVCOW7 (7.1 kb).

D. Construction of pGPF-IE-PA - Charts 5-6.

The assembly of pGPF-IE-PA is accomplished in two steps. First the GpF cDNA from pGPF3 is inserted into pSVCOW7 yielding pGPF-PA and then the immediate early promoter of cytomegalovirus is inserted to initiate transcription of the HRSV-like proteins yielding pGPF-IE-PA. STEP 1. Plasmid pSVCOW7 is cut with EcoRI and PuvI and fragment 7 (600 bp) containing the polyadenylation sequence of bovine growth hormone extending from the PvuII site in the 3' most exon of the BGH gene, to the EcoRI site downstream from the 3' end is isolated. For a complete discussion of the BGH polyadenylation sequence see the following references: (1) European patent application 0112012, published on 27 June 1984 wherein the identification and characterization of BGH genomic DNA is disclosed; (2) Woychik, R.P. et al., "Requirement for the 3' Flanking Region of the Bovine Growth Hormone Gene for Accurate Polyadenylation", Proc. Natl. Acad. Sci. USA 81:3944-3948 (July 1984); and, D.R. Higgs, et al., Nature 306:398-400 (24 November 1983) and references cited therein disclosing that the nucleotide sequence AATAAA characterizes the polyadenylation signal at a location 11 to 30 nucleotides upstream (towards the 5' end) from the 3' end of the BGH gene.

A second sample of pSVCOW7 is cut with EcoRI and BamHI to yield fragment 8. Fragment 8 can be alternatively derived from the EcoRI/BamHI fragment from parent plasmid pSV2dhfr available from

Bethesda Research Laboratories. Fragment 8 contains the origin of replication from pBR322 and an ampicillin resistance gene expressed in E. coli which allows for the selection of the plasmid in E. coli. The fragment also contains the mouse dihydrofolate reductase cDNA in a construction that allows expression in mammalian cells. Subramani, et al., Mol. Cell. Biol. 1:854-864 (1981).

Plasmid pGPF3 is cut with HindIII, treated with Klenow enzyme and recut with BamHI to yield fragment 9 (1.9 kb) which is gel isolated. Fragment 9 contains the leader and structural coding sequences from GpF cDNA. The BamHI site is just upstream from the cDNA coding for the 5' untranslated sequences of the mRNA, and the HindIII site is in pUC12 vector a few bases pairs beyond the PstI site near the 3' end of the gpF cDNA.

Fragments 7, 8 and 9 are ligated to form pGPF-PA (7.3 kb) which is a replication vector capable of shuttling between E coli and CHO cells. Plasmid pGPF-PA is transformed into E coli.

STEP 2. In step 2, pGPF-PA is converted into expression plasmid pGPF-IE-PA by inserting the immediate early gene promoter from human cytomegalovirus (CMV I.E. promoter). The CMV I.E. promoter is obtained from the PstI digestion of the CMV genome. The restriction endonuclease cleavage maps of the region of the human cytomegalovirus (CMV) genome containing the major immediate early gene (CMV I.E.) have been described in detail Stinski, et al., J. Virol. 46:1-14, 1983; Stenberg, et al., J. Virol. 49:190-199, 1984; and, Thomsen, et al., Proc. Natl. Acad. Sci. USA, 81:659-663, 1984.

The Stinski and Thomsen references describe a 2.0 kilobase PstI fragment which contains the promoter for the major immediate early gene. When this 2.0 kb PstI fragment is isolated and digested with Sau3AI, a 760 basepair fragment is obtained among the products. This 760 base pair fragment can be distinguished from the other products by its size and the presence of a SacI cleavage site and a BalI cleavage site within the fragment. Because of its convenient identification, utilization of this Sau3AI fragment is the preferred method of use of the CMV I.E. promoter as described in the present specification.

Plasmid pGPF-PA is cleaved with BamHI, and a Sau3AI fragment containing the CMV immediate early promoter is ligated into the compatible BamHI site. Plasmids containing the CMV promoter

fragment in an orientation such that transcription from the promoter would synthesize an mRNA for an HRSV-like protein are identified by cleavage of the plasmids with SacI. The resulting plasmid is designated pGPF-IE-PA having the CMV I.E. promoter at the 5'-end of the cDNA and the BGH polyadenylation signal on its 3'-end. The same procedures are used to obtain an equivalent expression vector for GpG. The plasmid is maintained in E. coli until transfection into CHO cells.

E. Transfection and Culturing of CHO Cells.

Plasmid pGPF-IE-PA is transfected into Chinese hamster ovary (CHO) cells deficient in dihydrofolate reductase(dhfr) using the calcium phosphate method for transfection of DNA into cells which is described in detail by Graham, et al. (in Introduction of Macromolecules into Viable Mammalian Cells, Alan R. Liss Inc., N.Y., 1980, pp. 3-25). The cell line used is the mutant DXB-11 originally available from L. Chasin, of Columbia University and completely described in Proc. Natl. Acad. Sci. USA 77:4216-4220, 1980). The above methods for transfection relies on the fact that cells which incorporate the transfected plasmids are no longer dhfr deficient and will grow in Dulbecco's modified Eagle's medium plus proline.

CHO cells expressing an HRSV-like protein are washed in phosphate buffered saline (PBS) at pH 7.4 and then lysed in PBS containing 1.0% Triton X-100 and 1.0% sodium deoxycholate. After pelleting the nuclei, the supernatant is applied to a conconavalin A column. The glycoproteins are eluted after extensive washing with a linear gradient of α-D-methylglucoside (0-0.5 M) in the above buffer. The eluted glycoproteins are dialyzed against PBS containing 0.1% Triton X-100 and applied to an affinity column. The affinity column is composed of either polyclonal or monoclonal antibodies of HRSV linked to Sepharose 4B beads (Pharmacia, Piscataway, New Jersey) by known techniques. The column is washed in dialysis buffer and the HRSV glycoproteins are eluted with PBS containing 0.1M glycine (pH 2.5) and 0.1% Triton X-100. The glycoprotein is dialyzed against saline and checked for purity by electrophoresis on a SDS-PAGE gel.

Example 3. The expression of HRSV GPF using Bovine Papilloma Virus (BPV)

A. The construction of a cloning vector containing a non-

transcribable expression cassette suitable for replication in E. coli.

The constructions of pTFW8 and pTFW9 offer a convenient starting material for expressing HRSV proteins using BPV. The transcription terminator of the deposited plasmid prevents the expression of HRSV proteins and must be removed in a single step excision and ligation.

a. Construction of PTFW8 - Chart 7.

Plasmid pdBPV-MMTneo (342-12) described in Mol. and Cell Biol., Vol 3 (No. 11):2110-2115 (1983) and obtained from Peter Howley of the National Cancer Institute, Bethesda, Maryland, USA. Plasmid pdBPV-MMT neo (342-12) consists of three parts: a complete BPV-1 genome (100%) opened at the unique BamHI site; pML2 (a "poison-minus" derivative of pBR322); and a transcriptional cassette composed of the murine metallothionein I gene promoter, the neomycin phosphotransferase II gene of Tn5, and the simian virus 40 early-region transcriptional processing signals. Plasmid pdBPV-MMT neo (342-12) is first digested with BamHI to remove the BPV sequences which were isolated and stored for later insertion. The remaining fragment is religated using T4 ligase to form pMMpro.nptII (6.7 kb). Removal of the BPV genome facilitates later genetic manipulations by creating unique restriction sites in the remaining plasmid. After the recombinations are complete, the BPV genome is replaced.

Plasmid pMMpro.nptII was digested with BglII and a synthetic DNA fragment 11 containing unique restriction sites is inserted and ligated using T4 ligase to yield pTFW8 (6.7 kb). Plasmid pTFW8 is identical to pMMpro.nptII except for the insertion of unique restriction sites between the murine metallothionein I gene promoter and the neomycin resistance gene.

b. Construction of pTWF9 - Chart 8.

Plasmid pTWF9 contains the transcription terminator T_I from phage lambda inserted between the metallothionein I gene promoter and the neomycin resistance gene. The transcription terminator can be obtained from Donald Court of the National Cancer Institute in Bethesda, Maryland USA. The transcription terminator is supplied in pKG1800sib3 which is the same as pUS6 as described in Gene, 28:343-350 (1984), except that t_I carries the sib3 mutation as described in guarneros et al., PNAS, 79:238-242 (1982). During the normal infection process of phage lambda, the t_I terminator functions in the

inhibition of bacteriophage λ int gene expression from P_L and in the termination of int gene transcription originating from P_I. The terminator is excised from pKG1800sib3 using AluI and PvuI as fragment 12 (1.2 kb), which is gel isolated and XhoI linkers are placed on either end of the fragment. The linkers are available from New England Biolabs, Beverly, MA, USA. The terminator fragment bounded by XhoI complementary ends is then inserted into pTWF8 which has been previously digested with XhoI. The fragements are then ligated using T4 DNA ligase to yield pTWF9 (7.9 kb). Plasmid pTWF9 was desposted in accordance with the Budapest Treaty. Plasmid pTFW9 is maintained in an E. coli host and has been deposited with the Northern Regional Research Center, Peoria, Illinois, USA on November 17, 1986 and assigned Accession Number NRRL B-18141.

B. The construction of pTFW/GPF - Chart 9.

In this example secretion of the glycoprotein into the culture media is desired. Therefore a universal translation termination oligonucleotide is ligated into the NsiI restriction enzyme site of the gpF gene in pGPF4 to cause a truncated glycoprotein which is missing its "anchor region" as described earlier. The modified plasmid is designated pGPF5. To construct pTFW/GPF, pGPF5 is digested with BamHI and HindIII. Its ends are made flush with Klenow enzyme and synthetic BglII linkers (New England Biolabs) are ligated to the ends of the clone. The DNA is digested with BglII and designated fragment 13 (1.9 kb). Fragment 13 containing the gpF gene is then isolated from a gel. The purified fragment is ligated into pTFW9 which has been digested with BglII to yield pTFW/GPF (9.8 kb).

C. Conversion of pTFW/GPF into a eukaryote expression vector-Chart 10.

Plasmid pTFW/GPF is converted into a eukaryote expression vector by reinserting the 100% complete BPV-1 genome excised with BamHI in step a., of Example 3A (Chart 7, step a). Plasmid pTFW/GPF is cut with BamHI and the BPV-1 intact genome, a 7.9 kb fragment (Chart 7), is inserted to yield pTFW/GPF/BPV* (17.7 kb) which is replicated in E. coli until production of glycoprotein F by eukaryotic cells is desired.

D. Expression of gpF in murine C127 cells.

Prior to transfection into murine C127 cells, pTFW/GPF/BPV* is digested with XhoI to excise the T_I terminator and religated with T4

DNA ligase (Chart 10). The resulting plasmid pTFW/GPF/BPV (16.5 kb) will now direct the expression of high levels of gpF which is secreted into the culture media. The C127 cells are available from the American Type Culture Collection and grown in Dulbecco's modified minimal essential media containing 10% fetal calf serum. The levels of gpF proteins in the media of the C127 cells are determined by Western blot experiments with anti-RSV antibody and 125_I-labeled protein A.

HRSV gpF (truncated) is purified by collecting the culture media surrounding the expressing cells. Serum-free media is preferred at this point if the levels of expression are acceptable in this media. The media is clarified by low speed centrifugation and concentrated by filtration. HRSV gpF is then purified by column chromatography as described for glycoproteins produced in CHO cells.

Example 4 The Expression of HRSV GPF Using Baculovirus Virus.

The following example relates to the expression of glycoprotein F in insect cell cultures. All procedures are detailed in Summers, M.D. and Smith, G.E., A Manual for Baculovirus Vectors and Insect Cell Culture Procedures published by the College of Agriculture, Texas Agricultural Experiment Station, Texas Agricultural Extension Service, College Station, Texas, 1986. The starting plasmid pAc373 (7.1 kb) is a general baculovirus expression vector having a unique BamHI site immediately downstream from the polyhedron promoter for Autographa californica nuclear polyhedrosis virus (AcNPV). The polyhedron protein is a matrix protein that is nonessential for viral infection and replication in vitro. The plasmid is available from Professor Max Summers of the Department of Entomology, Texas A & M University, College Station, Texas 77843 and is fully described in Molecular and Cell. Biology, 3(12):2156-2165 (1983).

A. Construction of pAcGPF - Chart 11.

Plasmid pGPF5 is digested with HindIII and the ends are made flush with Klenow enzymme. Synthetic BamHI linkers (New England Biolabs) are ligated to the end of the DNA. The DNA is digested with BamHI and fragment 14 containing the gpF gene is isolated from a gel. The purified fragment is ligated into pAc373 which has been digested with BamHI.

B. Transfection and culturing of S. Frugiperda.

The gpF cDNA insert of pAcGPF is recombined with native AcNPV DNA by cotransfection in S. frugiperda. S. Frugiperda (SF9; ATCC CRL 1711) are cultured in Grace Media (Gibco Lab. Livonia, MI 48150), 10% fetal calf serum and supplemented with Difco Lactalbumin hydrolysolate and yestolate. The cells are cotransfected with AcNPV DNA and pAcGPF at 1μ/ml and 2μ/ml respectively. Resulting virus particles are obtained by collecting the media and removing cellular material by low speed centrifugation. The virus containing-media is then used to infect S. frugiperda. Subsequent infection of S. frugiperda using these viral particles which include both native viral DNA and DNA recombined with the cDNA coding for glycoprotein F will result in some cells expressing the HRSV protein instead of the polyhedron protein. Purification of recombinant virus is accomplished by a series of limited dilution platings in 96-well tissue culture plates containing S. frugiperda cells. Wells containing recombinant virus are deteted by dot blot hybridization using pGPF4 which has been labeled with ^{32}p-dCTP by nick translation as a probe. Once sufficiently pure, the recombinant virus is detected by its unique inclusion-negative plaque morphology. HRSV protein synthesized in recombinant baculovirus infected cells is detected by Western blot experiments with anti-RSV antibody and ^{125}I-labeled protein A (Amersham Corp.).

The HRSV protein is purified from the culture media by the methodology described in the BPV expression system for C125 cells.

Example 5 Preparation of a Vaccine for HRSV

The immunogen can be prepared in vaccine dose form by well-known procedures. The vaccine can be administered intramuscularly, subcutaneously or intranasally. For parenteral administration, such as intramuscular injection, the immunogen may be combined with a suitable carrier, for example, it may be administered in water, saline or buffered vehicles with or without various adjuvants or immunomodulating agents such as aluminum hydroxide, aluminum phosphate, aluminum potassium sulfate (alum), beryllium sulfate, silica, kaolin, carbon, water-in-oil emulsions, oil-in-water emulsions, muramyl dipeptide, bacterial endotoxin, lipid X, Corynebacterium parvum (Propionobacterium acnes), Bordetella pertussis, polyribonucleotides, sodium alginate, lanolin, lysolecithin, vitamin A, saponin, liposomes, levamisole, DEAE-dextran, blocked copolymers or

other synthetic adjuvants. Such adjuvants are available commercially from various sources, for example, Merck Adjuvant 65 (Merck and Company, Inc., Rahway, N.J.).

The proportion of immunogen and adjuvant can be varied over a broad range so long as both are present in effective amounts. For example, aluminum hydroxide can be present in an amount of about 0.5% of the vaccine mixture (Al_2O_3 basis). On a per dose basis, the concentration of the immunogen can range from about 0.015 μg to about 1.5 mg per kilogram per patient body weight. A preferable dosage range is from about 1.5 μg/kg to about .043 mg/kg of patient body weight. A suitable dose size in humans is about 0.1 - 1 ml, preferably about 0.1 ml. Accordingly, a dose for intramuscular injection, for example, would comprise 0.1 ml containing immunogen in admixture with 0.5% aluminum hydroxide.

The vaccine can be administered to pregnant women or to women of child-bearing age to stimulate maternal HRSV antibodies. The female can be revaccinated as needed. Infants can be vaccinated at 2 to 3 months of age and revaccinated as necessary, preferably at 6 to 9 months of age. Babies born to unvaccinated mothers can be vaccinated at 2 to 3 months of age. The vaccine may also be useful in other susceptible populations such as elderly or infirmed patients.

The vaccine may also be combined with other vaccines for other diseases to produce multivalent vaccines. It may also be combined with other medicaments such as antibiotics.

CHART 1. CONSTRUCTION OF pGPF

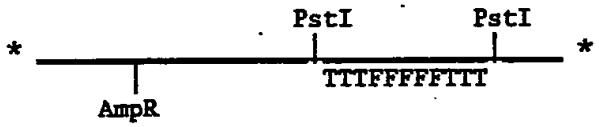
(a) Plasmid pBR322 is cut with PstI and tailed with guanosine to yield fragment 1 which is gel isolated.

Fragment 1

(b) cDNA from mRNA of HRSV is tailed with 10-15 dCMP residues per 3' end.

(c) Fragment 1 and the cDNA from HRSV mRNA are ligated and pGPF identified by hybridization with the appropriate probe.

pGPF



AmpR = Ampicillin resistance.
T = Guanosine/cytosine tail.
F = Glycoprotein F.

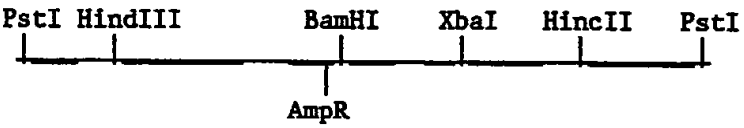
CHART 2. CONSTRUCTION OF pGPF2

(a) Plasmid pGPF is cut with PstI and fragment 2 (1.9 kb) is gel isolated.

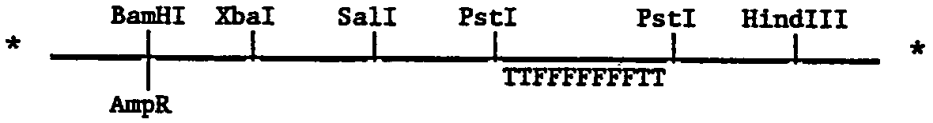
Fragment 2

(b) Plasmid pUC12 (2.7 kb) is cut with PstI to yield fragment 3 which is gel isolated.

Fragment 3



(c) Fragments 4 and 5 are ligated to yield pGPF2 (4.6 kb) which is transformed in E. coli.

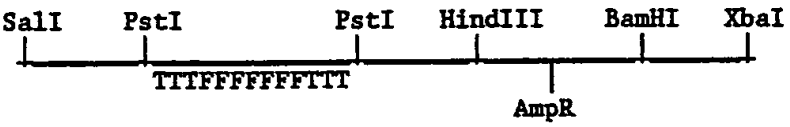


AmpR = Ampicillin resistance.
T = Guanosine/cytosine tail.
F = Glycoprotein F.

CHART 3. CONSTRUCTION OF pGPF3 AND pGPF4

(a) Plasmid pGPF2 is cut with XbaI, treated with bacterial alkaline phosphatase, recut with SalI and treated with Klenow enzyme to yield fragment 4.

Fragment 4

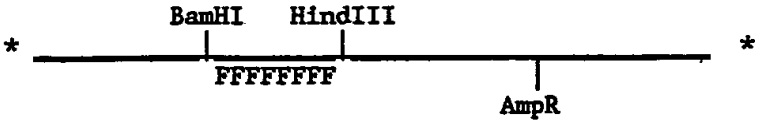


(b) Fragment 4 is digested downstream from the SalI site using lambda exonuclease and the remaining 3' tail is hybridized to the synthetic oligonucleotide complementary to the 5' portion of the leader sequence having the following sequence of GpF cDNA.

5'-end ATGGAGTTGCTAATC

(c) The single stranded portion of the cDNA 3' downstream from the synthetic oligonucleotides are filled in using Klenow enzyme and the ends are ligated using T4 ligase to yield pGPF3 (4.6 kb).

(d) Plasmid pGPF3 is cut with HindIII and treated with Bal 31 to digest the G-C nucleotide tail at the 3' end of the gpF CDNA. The gpF cDNA is cut with BamHI (1.7 kb) isolated from a gel and religated into a BamHI/HincII digestion of PUC12 to yield pGPF4 (4.4 kb).

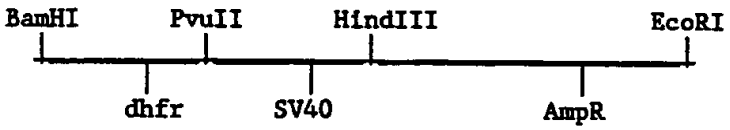


AmpR = Ampicillin resistance.
T = Guanosine/cytosine tail.
F = Glycoprotein F.

CHART 4. CONSTRUCTION OF pSVCOW7

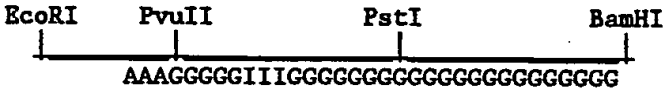
(a) Plasmid pSV2dhfr is cut with BamHI and EcoRI to obtain fragment 5 (5.0 kb).

Fragment 5



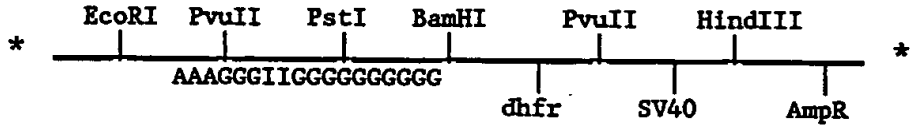
(b) Plasmid pλGH2R2 is cut with BamHI and EcoRI to obtain fragment 6 (2.1 kb).

Fragment 6



(c) Fragments 5 and 6 are ligated to yield pSVCOW7 (7.1 kb).

pSVCOW7



A = Bovine growth hormone poly A tail.
G = Genomic bovine growth hormone.
I = Intron.
dhfr = Dihydrofolate reductase.
SV40 = SV40 promoter and origin of replication.
AmpR = Ampicillin resistance

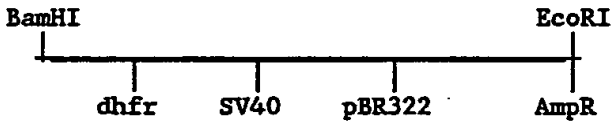
CHART 5. CONSTRUCTION OF pGPF-PA

(a) pSVCOW7 is cut with EcoRI and PvuII to yield fragment 7 (600 bp) containing the polyadenylation sequence of bovine growth hormone which is gel isolated.

Fragment 7

(b) pSVCOW7 is cut with EcoRI and BamHI to yield fragment 8 (5.8 kb).

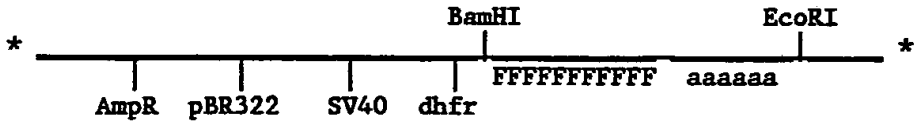
Fragment 8



(c) Plasmid pGPF4 is cut with HindIII, treated with Klenow enzyme, cut with BamHI to yield fragment 9 (1.9 kb) containing GPF having a 3' BamHI overhang upstream and a blunt end downstream from the message.

Fragment 9

(d) Fragments 5, 6 and 7 are ligated to form pGPF-PA which is maintained in E. coli.



AmpR = Ampicillin resistance.
pBR322 = Replication origin for pBR322
SV40 = Replication origin for SV40
dhfr = dihydrofolate reductase
F = Glycoprotein F.
CMV = Cytomegalovirus promoter.
a = Polyadenylation tail.
T = guanosine/cytosine tail

CHART 6. CONSTRUCTION OF pGPF-IE-PA

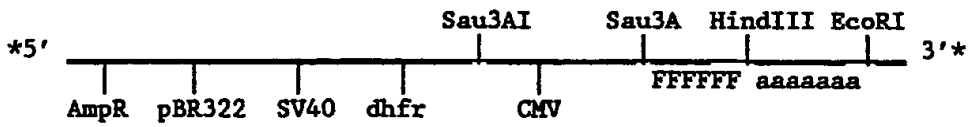
(a) Plasmid pGPF-PA is cut with BamHI to yield fragment 10 (7.3 kb).

Fragment 10

(b) The CMV immediate early promoter is obtained from a Sau3AI digestion of a PstI fragment from the CMV genome. Sau3A is compatible with BamHI for ligation.

(c) Fragment 10 and the CMV promoter are ligated to yield pGPF-IE-PA (8.0 kb).

pGPF-IE-PA

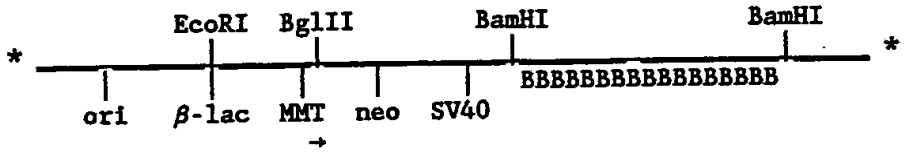


AmpR = Ampicillin resistance.
pBR322 = Replication origin for pBR322
SV40 = Replication origin for SV40
dhfr = dihydrofolate reductase
F = Glycoprotein F.
CMV = Cytomegalovirus promoter.
a = Polyadenylation tail.

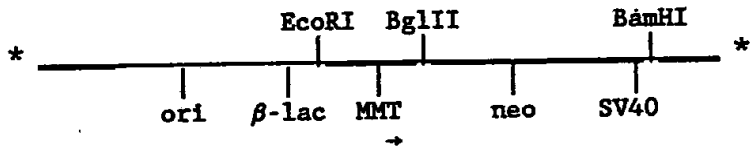
CHART 7. Construction of pTFW8

a) Plasmid pdBPV-MMTneo (342-12) (14.6 kb) was cut with BamHI and the bovine papilloma virus genome was excised (7.9 kb) gel isolated and saved. The remaining fragment was gel isolated, religated using T4 ligase and designated pMMpro.nptII (6.7 kb).

pdBPV-MMTneo (342-12)

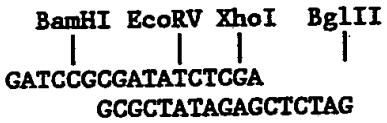


pMMpro.nptII

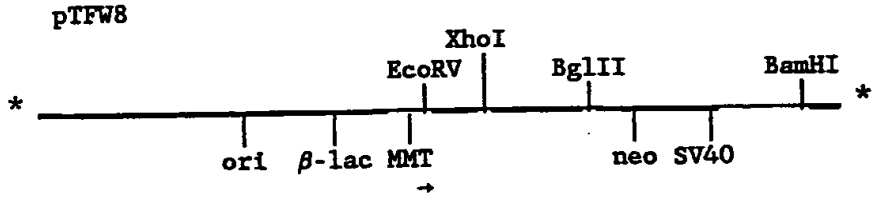


b) Plasmid pMMpro.nptII was cut with BglII and synthetic fragment 11 inserted and the plasmid religated to yield pTFW8 (6.7 kb).

Synthetic Fragment 11



pTFW8



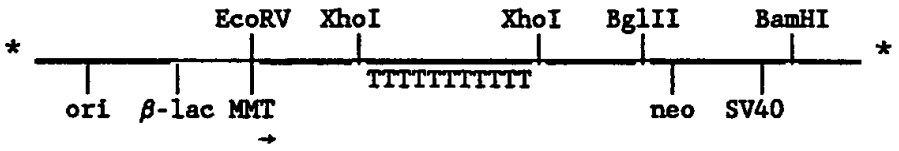
B = Bovine papilloma virus sequences.
SV40 = Simian virus 40 sequences early region, small t antigen splicing signals and 3' transcriptional processing signals.
neo = Neomycin phosphotransferase II gene (NPTII).
MMT = Murine metallothionein gene promoter.
ori = pBR322 origin of replication.
β-Lac = β-lactamase gene.

CHART 8. Construction of pTFW9

a) Plasmid pTFW8 (Chart 1) is cut with XhoI and fragment 12 containing the t_I terminator from pKG1800sib3 is inserted using T4 ligase to obtain plasmid pTFW9 (7.9 kb).

Fragment 12

pTFW9

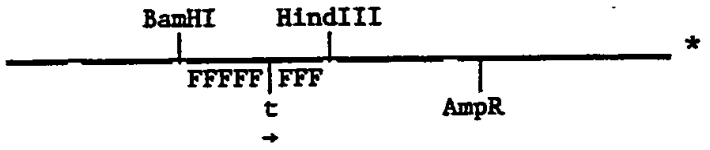


SV40 - Simian virus 40 sequences early region, small t antigen splicing signals and 3' transcriptional processing signals.
neo - Neomycin phosphotransferase II gene (NPTII).
MMT - Murine metallothionein gene promoter.
ori - pBR322 origin of replication.
β-Lac - β-lactamase gene.
T - λt_I terminator.

CHART 9. Construction of pTFW/GPF

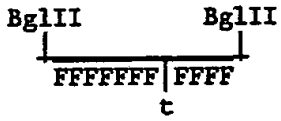
a) pGPF4 (Chart 3) is cut with NsiI and a translation terminator ligated into the CDNA of gpF yielding pGPF5 (4.6 kb).

pGPF5

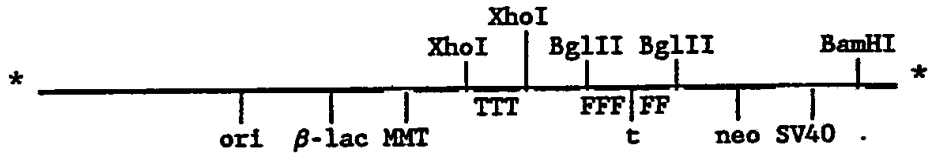


b) Plasmid pGPF5 is cut with BamHI and HindIII isolating Fragment 13 consisting of the cDNA encoding gpF (1.9 kb). The ends of fragment 13 are made blunt with Klenow enzyme and synthetic BgIII linkers are ligated to the ends of the clone and the cDNA treated with BglII to yield Fragement 13.

Fragement 13



c) Plasmid pTFW9 is cut with BglII and Fragment 13 is inserted and religated to form pTFW/GPF (9.8 kb).

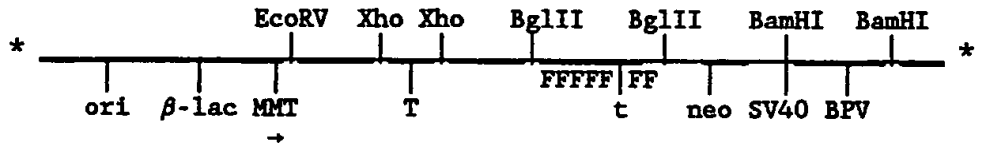


SV40 - Simian virus 40 sequences early region, small t antigen splicing signals and 3' transcriptional processing signals.
neo - Neomycin phosphotransferase II gene (NPTII).
MMT - Murine metallothionein gene promoter.
ori - pBR322 origin of replication.
β-Lac - β-lactamase gene
F - glycoprotein F
t - translation terminator.
AmpR - Ampicillin resistance.

CHART 10. Construction of pTFW/GPF/BPV

a) Plasmid pTFW/GPF (Chart 9) is cut with BamHI and the intact BPV genome (from chart 7 step a) is inserted and ligated into pTFW/GPF to yield pTFW/GPF/BPV* (17.9 kb).

pTFW/GPF/BPV*



b) pTFW9/GPF/BPV* is cut with Xho and the large fragment is religated to yield pTFW/GPF/BPV (9.3 kb).

pTFW/GPF/BPV

SV40 - Simian virus 40 sequences early region, small t antigen splicing signals and 3' transcriptional processing signals.
neo - Neomycin phosphotransferase II gene (NPTII).
MMT - Murine metallothionein gene promoter.
ori - pBR322 origin of replication.
β-Lac - β-lactamase gene
F - gpF protein
t - translation terminator.
T - λt_I terminator.

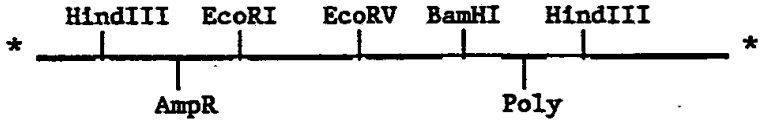
CHART 11. Construction of pAcGPF

a) Plasmid pGPF5 (chart 9) is cut with HindIII and the ends made flush with Klenow enzyme. Synthetic BamHI linkers are ligated and the plasmid digested with BamHI to yield fragment 14 (1.9 kb) containing the gpF cDNA. Fragment 14 is gel isolated.

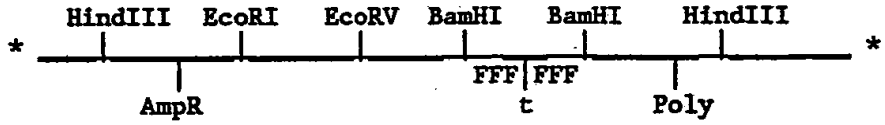
Fragment 14

b) pAc373 (7.1 kb) is treated with BamHI to linearize pAc373 (7.1 kb).

pAc373



c) The linear pAc373 and fragment 14 are annealed and ligated to form pAcGPF (9.0 kb).



N = Untranslated 3' portion of TPAcDNA.
AmpR = Ampicillin resistance.
Poly = Polyhedrin protein gene.
F = glycoprotein F.
t = translation terminator.

CHART 12

Nucleotide sequence of the F mRNA
and the predicted protein sequence

```
                                G GGG CAA ATA ACA ATG GAG TTG CTA ATC CTC AAA     34
                                              Met Glu Leu Leu Ile Leu Lys          7

GCA AAT GCA ATT ACC ACA ATC CTC ACT GCA GTC ACA TTT TGT TTT GCT TCT GGT          88
Ala Asn Ala Ile Thr Thr Ile Leu Thr Ala Val Thr Phe Cys Phe Ala Ser Gly         25

CAA AAC ATC ACT GAA GAA TTT TAT CAA TCA ACA TGC AGT GCA GTT AGC AAA GGC         142
Gln Asn Ile Thr Glu Glu Phe Tyr Gln Ser Thr Cys Ser Ala Val Ser Lys Gly         43

TAT CTT AGT GCT CTG AGA ACT GGT TGG TAT ACC AGT GTT ATA ACT ATA GAA TTA         196
Tyr Leu Ser Ala Leu Arg Thr Gly Trp Tyr Thr Ser Val Ile Thr Ile Glu Leu         61

AGT AAT ATC AAG GAA AAT AAG TGT AAT GGA ACA GAT GCT AAG GTA AAA TTG ATA         250
Ser Asn Ile Lys Glu Asn Lys Cys Asn Gly Thr Asp Ala Lys Val Lys Leu Ile         79

AAA CAA GAA TTA GAT AAA TAT AAA AAT GCT GTA ACA GAA TTG CAG TTG CTC ATG         304
Lys Gln Glu Leu Asp Lys Tyr Lys Asn Ala Val Thr Glu Leu Gln Leu Leu Met         97

CAA AGC ACA CCA CCA ACA AAC AAT CGA GCC AGA AGA GAA CTA CCA AGG TTT ATG         358
Gln Ser Thr Pro Pro Thr Asn Asn Arg Ala Arg Arg Glu Leu Pro Arg Phe Met        115

AAT TAT ACA CTC AAC AAT GCC AAA AAA ACC AAT GTA ACA TTA AGC AAG AAA AGG         412
Asn Tyr Thr Leu Asn Asn Ala Lys Lys Thr Asn Val Thr Leu Ser Lys Lys Arg        133

AAA AGA AGA TTT CTT GGT TTT TTG TTA GGT GTT GGA TCT GCA ATC GCC AGT GGC         466
Lys Arg Arg Phe Leu Gly Phe Leu Leu Gly Val Gly Ser Ala Ile Ala Ser Gly        151

GTT GCT GTA TCT AAG GTC CTG CAC CTA GAA GGG GAA GTG AAC AAG ATC AAA AGT         520
Val Ala Val Ser Lys Val Leu His Leu Glu Gly Glu Val Asn Lys Ile Lys Ser        169
```

CHART 12 (Continued)

```
GCT CTA CTA TCC ACA AAC AAG GCT GTA GTC AGC TTA TCA AAT GGA GTT AGT GTC   574
Ala Leu Leu Ser Thr Asn Lys Ala Val Val Ser Leu Ser Asn Gly Val Ser Val   187

TTA ACC AGC AAA GTG TTA GAC CTC AAA AAC TAT ATA GAT AAA CAA TTG TTA CCT   628
Leu Thr Ser Lys Val Leu Asp Leu Lys Asn Tyr Ile Asp Lys Gln Leu Leu Pro   205

ATT GTG AAC AAG CAA AGC TGC AGC ATA TCA AAT ATA GAA ACT GTG ATA GAG TTC   682
Ile Val Asn Lys Gln Ser Cys Ser Ile Ser Asn Ile Glu Thr Val Ile Glu Phe   223

CAA CAA AAG AAC AAC AGA CTA CTA GAG ATT ACC AGG GAA TTT AGT GTT AAT GCA   736
Gln Gln Lys Asn Asn Arg Leu Leu Glu Ile Thr Arg Glu Phe Ser Val Asn Ala   241

GGT GTA ACT ACA CCT GTA AGC ACT TAC ATG TTA ACT AAT AGT GAA TTA TTG TCA   790
Gly Val Thr Thr Pro Val Ser Thr Tyr Met Leu Thr Asn Ser Glu Leu Leu Ser   259

TTA ATC AAT GAT ATG CCT ATA ACA AAT GAT CAG AAA AAG TTA ATG TCC AAC AAT   844
Leu Ile Asn Asp Met Pro Ile Thr Asn Asp Gln Lys Lys Leu Met Ser Asn Asn   277

GTT CAA ATA GTT AGA CAG CAA AGT TAC TCT ATC ATG TCC ATA ATA AAA GAG GAA   898
Val Gln Ile Val Arg Gln Gln Ser Tyr Ser Ile Met Ser Ile Ile Lys Glu Glu   295

GTC TTA GCA TAT GTA GTA CAA TTA CCA CTA TAT GGT GTT ATA GAT ACA CCC TGT   952
Val Leu Ala Tyr Val Val Gln Leu Pro Leu Tyr Gly Val Ile Asp Thr Pro Cys   313

TGG AAA CTA CAC ACA TCC CCT CTA TGT ACA ACC AAC ACA AAA GAA GGG TCC AAC  1006
Trp Lys Leu His Thr Ser Pro Leu Cys Thr Thr Asn Thr Lys Glu Gly Ser Asn   331

ATC TGT TTA ACA AGA ACT GAC AGA GGA TGG TAC TGT GAC AAT GCA GGA TCA GTA  1060
Ile Cyc Leu Thr Arg Thr Asp Arg Gly Trp Tyr Cys Asp Asn Ala Gly Ser Val   349

TCT TTC TTC CCA CAA GCT GAA ACA TGT AAA GTT CAA TCA AAT CGA GTA TTT TGT  1114
Ser Phe Phe Pro Gln Ala Glu Thr Cys Lys Val Gln Ser Asn Arg Val Phe Cys   367
```

CHART 12 (Continued)

```
GAC ACA ATG AAC AGT TTA ACA TTA CCA AGT GAA ATA AAT CTC TGC AAT GTT GAC   1168
Asp Thr Met Asn Ser Leu Thr Leu Pro Ser Glu Ile Asn Leu Cys Asn Val Asp    385

ATA TTC AAC CCC AAA TAT GAT TGT AAA ATT ATG ACT TCA AAA ACA GAT GTA AGC   1222
Ile Phe Asn Pro Lys Tyr Asp Cys Lys Ile Met Thr Ser Lys Thr Asp Val Ser    403

AGC TCC GTT ATC ACA TCT CTA GGA GCC ATT GTG TCA TGC TAT GGC AAA ACT AAA   1276
Ser Ser Val Ile Thr Ser Leu Gly Ala Ile Val Ser Cys Tyr Gly Lys Thr Lys    421

TGT ACA GCA TCC AAT AAA AAT CGT GGA ATC ATA AAG ACA TTT TCT AAC GGG TGC   1330
Cys Thr Ala Ser Asn Lys Asn Arg Gly Ile Ile Lys Thr Phe Ser Asn Gly Cys    439

GAT TAT GTA TCA AAT AAA GGG ATG GAC ACT GTG TCT GTA GGT AAC ACA TTA TAT   1384
Asp Tyr Val Ser Asn Lys Gly Met Asp Thr Val Ser Val Gly Asn Thr Leu Tyr    457

TAT GTA AAT AAG CAA GAA GGT AAA AGT CTC TAT GTA AAA GGT GAA CCA ATA ATA   1438
Tyr Val Asn Lys Gln Glu Gly Lys Ser Leu Tyr Val Lys Gly Glu Pro Ile Ile    475

AAT TTC TAT GAC CCA TTA GTA TTC CCC TCT GAT GAA TTT GAT GCA TCA ATA TCT   1492
Asn Phe Tyr Asp Pro Leu Val Phe Pro Ser Asp Glu Phe Asp Ala Ser Ile Ser    493

CAA GTC AAC GAG AAG ATT AAC CAG AGC CTA GCA TTT ATT CGT AAA TCC GAT GAA   1546
Gln Val Asn Glu Lys Ile Asn Gln Ser Leu Ala Phe Ile Arg Lys Ser Asp Glu    511

TTA TTA CAT AAT GTA AAT GCT GGT AAA TCC ACC ACA AAT ATC ATG ATA ACT ACT   1600
Leu Leu His Asn Val Asn Ala Gly Lys Ser Thr Thr Asn Ile Met Ile Thr Thr    529

ATA ATT ATA GTG ATT ATA GTA ATA TTG TTA TCA TTA ATT GCT GTT GGA CTG CTC   1654
Ile Ile Ile Val Ile Ile Val Ile Leu Leu Ser Leu Ile Ala Val Gly Leu Leu    547

TTA TAC TGT AAG GCC AGA AGC ACA CCA GTC ACA CTA AGC AAA GAT CAA CTG AGT   1708
Leu Tyr Cys Lys Ala Arg Ser Thr Pro Val Thr Leu Ser Lys Asp Gln Leu Ser    565
```

CHART 12 (Continued)

```
GGT ATA AAT AAT ATT GCA TTT AGT AAC TAA ATA AAA ATA GCA CCT AAT CAT GTT   1762
Gly Ile Asn Asn Ile Ala Phe Ser Asn                                      574

CTT ACA ATG GTT TAC TAT CTG CTC ATA GAC AAC CCA TCT GTC ATT GGA TTT TCT   1816

TAA AAT CTG AAC TTC ATC GAA ACT CTC ATC TAT AAA CCA TCT CAC TTA CAC TAT   1870

TTA AGT AGA TTC CTA GTT TAT AGT TAT AT                                    1879
```

CHART 13

Nucleotide sequence of the RS virus G mRNA
and the predicted protein sequence

```
1    GGG GCA AAT GCA AAC ATG TCC AAA AAC AAG GAC CAA CGC ACC GCT AAG ACA TTA
                        Met Ser Lys Asn Lys Asp Gln Arg Thr Ala Lys Thr Leu

55   GAA AGG ACC TGG GAC ACT CTC AAT CAT TTA TTA TTC ATA TCA TCG TGC TTA TAT
     Glu Arg Thr Trp Asp Thr Leu Asn His Leu Leu Phe Ile Ser Ser Cys Leu Tyr

109  AAG TTA AAT CTT AAA TCT GTA GCA CAA ATC ACA TTA TCC ATT CTG GCA ATG ATA
     Lys Leu Asn Leu Lys Ser Val Ala Gln Ile Thr Leu Ser Ile Leu Ala Met Ile

163  ATC TCA ACT TCA CTT ATA ATT GCA GCC ATC ATA TTC ATA GCC TCG GCA AAC CAC
     Ile Ser Thr Ser Leu Ile Ile Ala Ala Ile Ile Phe Ile Ala Ser Ala Asn His

217  AAA GTC ACA CCA ACA ACT GCA ATC ATA CAA GAT GCA ACA AGC CAG ATC AAG AAC
     Lys Val Thr Pro Thr Thr Ala Ile Ile Gln Asp Ala Thr Ser Gln Ile Lys Asn

271  ACA ACC CCA ACA TAC CTC ACC CAG AAT CCT CAG CTT GGA ATC AGT CCC TCT AAT
     Thr Thr Pro Thr Tyr Leu Thr Gln Asn Pro Gln Leu Gly Ile Ser Pro Ser Asn

325  CCG TCT GAA ATT ACA TCA CAA ATC ACC ACC ATA CTA GCT TCA ACA ACA CCA GGA
     Pro Ser Glu Ile Thr Ser Gln Ile Thr Thr Ile Leu Ala Ser Thr Thr Pro Gly

379  GTC AAG TCA ACC CTG CAA TCC ACA ACA GTC AAG ACC AAA AAC ACA ACA ACA ACT
     Val Lys Ser Thr Leu Gln Ser Thr Thr Val Lys Thr Lys Asn Thr Thr Thr Thr

433  CAA ACA CAA CCC AGC AAG CCC ACC ACA AAA CAA CGC CAA AAC AAA CCA CCA AGC
     Gln Thr Gln Pro Ser Lys Pro Thr Thr Lys Gln Arg Gln Asn Lys Pro Pro Ser

487  AAA CCC AAT AAT GAT TTT CAC TTT GAA GTG TTC AAC TTT GTA CCC TGC AGC ATA
     Lys Pro Asn Asn Asp Phe His Phe Glu Val Phe Asn Phe Val Pro Cys Ser Ile
```

CHART 13 (Continued)

```
541   TGC AGC AAC AAT CCA ACC TGC TGG GCT ATC TGC AAA AGA ATA CCA AAC AAA AAA
      Cys Ser Asn Asn Pro Thr Cys Trp Ala Ile Cys Lys Arg Ile Pro Asn Lys Lys

595   CCA GGA AAG AAA ACC ACT ACC AAG CCC ACA AAA AAA CCA ACC CTC AAG ACA ACC
      Pro Gly Lys Lys Thr Thr Thr Lys Pro Thr Lys Lys Pro Thr Leu Lys Thr Thr

649   AAA AAA GAT CCC AAA CCT CAA ACC ACT AAA TCA AAG GAA GTA CCC ACC ACC AAG
      Lys Lys Asp Pro Lys Pro Gln Thr Thr Lys Ser Lys Glu Val Pro Thr Thr Lys

703   CCC ACA GAA GAG CCA ACC ATC AAC ACC ACC AAA ACA AAC ATC ATA ACT ACA CTA
      Pro Thr Glu Glu Pro Thr Ile Asn Thr Thr Lys Thr Asn Ile Ile Thr Thr Leu

757   CTC ACC TCC AAC ACC ACA GGA AAT CCA GAA CTC ACA AGT CAA ATG GAA ACC TTC
      Leu Thr Ser Asn Thr Thr Gly Asn Pro Glu Leu Thr Ser Gln Met Glu Thr Phe

811   CAC TCA ACT TCC TCC GAA GGC AAT CCA AGC CCT TCT CAA GTC TCT ACA ACA TCC
      His Ser Thr Ser Ser Glu Gly Asn Pro Ser Pro Ser Gln Val Ser Thr Thr Ser

865   GAG TAC CCA TCA CAA CCT TCA TCT CCA CCC AAC ACA CCA CGC CAG TAG TTA CTT
      Glu Tyr Pro Ser Gln Pro Ser Ser Pro Pro Asn Thr Pro Arg Gln End

919   AAA AAA AAA AAA AAA AA 935
```

Complete nucleotide sequence of 22K mRNA and the predicted protein sequence encoded by the 5'-proximal open reading frame

```
GGG GCA AAT ATG TCA CGA AGG AAT CCT TGC AAA TTT GAA ATT CGA GGT CAT TGC TTA AAT GGT AAG AGG TGT CAT 75
      Met Ser Arg Arg Asn Pro Cys Lys Phe Glu Ile Arg Gly His Cyc Leu Asn Gly Lys Arg Cys His 22

TTT AGT CAT AAT TAT TTT GAA TGG CCA CCC CAT GCA CTG CTT GTA AGA CAA AAC TTT ATG TTA AAC AGA ATA CTT 150
Phe Ser His Asn Tyr Phe Glu Trp Pro Pro His Ala Leu Leu Val Arg Gln Asn Phe Met Leu Asn Arg Ile Leu 47

AAG TCT ATG GAT AAA AGT ATA GAT ACC TTA TCA GAA ATA AGT GGA GCT GCA GAG TTG GAC AGA ACA GAA GAG TAT 225
Lys Ser Met Asp Lys Ser Ile Asp Thr Leu Ser Glu Ile Ser Gly Ala Ala Glu Leu Asp Arg Thr Glu Glu Tyr 72

GCT CTT GGT GTA GTT GGA GTG CTA GAG AGT TAT ATA GGA TCA ATA AAC AAT ATA ACT AAA CAA TCA GCA TGT GTT 300
Ala Leu Gly Val Val Gly Val Leu Glu Ser Tyr Ile Gly Ser Ile Asn Asn Ile Thr Lys Gln Ser Ala Cys Val 97

GCC ATG AGC AAA CTC CTC ACT GAA CTC AAT AGT GAT GAT ATC AAA AAG CTG AGG GAC AAT GAA GAG CTA AAT TCA 375
Ala Met Ser Lys Leu Leu Thr Glu Leu Asn Ser Asp Asp Ile Lys Lys Leu Arg Asp Asn Glu Glu Leu Asn Ser 122

CCC AAG ATA AGA GTG TAC AAT ACT GTC ATA TCA TAT ATT GAA AGC AAC AGG AAA AAC AAT AAA CAA ACT ATC CAT 450
Pro Lys Ile Arg Val Tyr Asn Thr Val Ile Ser Tyr Ile Glu Ser Asn Arg Lys Asn Asn Lys Gln Thr Ile His 147

CTG TTA AAA AGA TTG CCA GCA GAC GTA TTG AAG AAA ACC ATC AAA AAC ACA TTG GAT ATC CAT AAG AGC ATA ACC 525
Leu Leu Lys Arg Leu Pro Ala Asp Val Leu Lys Lys Thr Ile Lys Asn Thr Leu Asp Ile His Lys Ser Ile Thr 172

ATC AAC AAC CCA AAA GAA TCA ACT GTT AGT GAT ACA AAT GAC CAT GCC AAA AAT AAT GAT ACT ACC TGA CAA ATA 600
Ile Asn Asn Pro Lys Glu Ser Thr Val Ser Asp Thr Asn Asp His Ala Lys Asn Asn Asp Thr Thr          194

TCC TTG TAG TAT AAC TTC CAT ACT AAT AAC AAG TAG ATG TAG ACT TAC TAT CTA TAA TCA AAA GAA CAC ACT ATA 675

TTT CAA TCA AAA CAA CCC AAA TAA CCA TAT GTA CTC ACC GAA TCA AAC ATT CAA TGA AAT CCA TTG GAC CTC TCA 750

AGA ATT GAT TGA CAC AAT TCA AAA TTT TCT ACA ACA TCT AGG TAT TAT TGA GGA TAT ATA TAC AAT ATA TAT ATT 825

AGT GTC ATA ACA CTC AAT TCT AAC ACT CAC CAC ATC GTT ACA TTA TTA ATT CAA ACA ATT CAA GTT GTG GGA CAA 900

AAT GGA TCC CAT TAT TAA TGG AAA TTC TGC TAA TGT TTA TCT AAC CGA TAG TTA TTT                        957
```

CHART 15

Complete nucleotide sequence of the 1A mRNA
and the predicted amino acid sequence of the encoded protein

```
GGG GCA AAT AAT CAT TGG AGG AAA TCC AAC TAA TCA CAA TAT CTG TTA ACA TAG ACA AGT 60

CCA CAC ACC ATA CAG AAT CAA CCA ATG GAA AAT ACA TCC ATA ACA ATA GAA TTC TCA AGC 120
                               Met Glu Asn Thr Ser Ile Thr Ile Glu Phe Ser Ser 12

AAA TTC TGG CCT TAC TTT ACA CTA ATA CAC ATG ATC ACA ACA ATA ATC TCT TTG CTA ATC 180
Lys Phe Trp Pro Tyr Phe Thr Leu Ile His Met Ile Thr Thr Ile Ile Ser Leu Leu Ile 32

ATA ATC TCC ATC ATG ATT GCA ATA CTA AAC AAA CTT TGT GAA TAT AAC GTA TTC CAT AAC 240
Ile Ile Ser Ile Met Ile Ala Ile Leu Asn Lys Leu Cys Glu Tyr Asn Val Phe His Asn 52

AAA ACC TTT GAG TTA CCA AGA GCT CGA GTC AAC ACA TAG CAT TCA TCA ATC CAA CAG CCC 300
Lys Thr Phe Glu Leu Pro Arg Ala Arg Val Asn Thr                                 64

AAA ACA GTA ACC TTG CAT TTA AAA ATG AAC AAC CCC TAC CTC TTT ACA ACA CCT CAT TAA 360

CAT CCC ACC ATG CAA ACC ACT ATC CAT ACT ATA AAG TAG TTA ATT                     405
```

Complete nucleotide sequence of the major nucleocapsid protein mRNA
and the predicted amino acid sequence

```
GGG GCA AAT ACA AAG ATG GCT CTT AGC AAA GTC AAG TTG AAT GAT ACA CTC AAC AAA GAT CAA CTT CTG TCA TCC 75
                Met Ala Leu Ser Lys Val Lys Leu Asn Asp Thr Leu Asn Lys Asp Gln Leu Leu Ser Ser 20

AGC AAA TAC ACC ATC CAA CGG AGC ACA GGA GAT AGT ATT GAT ACT CCT AAT TAT GAT GTG CAG AAA CAC ATC AAT 150
Ser Lys Tyr Thr Ile Gln Arg Ser Thr Gly Asp Ser Ile Asp Thr Pro Asn Tyr Asp Val Gln Lys His Ile Asn 45

AAG TTA TGT GGC ATG TTA TTA ATC ACA GAA GAT GCT AAT CAT AAA TTC ACT GGG TTA ATA GGT ATG TTA TAT GCG 225
Lys Leu Cys Gly Met Leu Leu Ile Thr Glu Asp Ala Asn His Lys Phe Thr Gly Leu Ile Gly Met Leu Tyr Ala 70

ATG TCT AGG TTA GGA AGA GAA GAC ACC ATA AAA ATA CTC AGA GAT GCG GGA TAT CAT GTA AAA GCA AAT GGA GTA 300
Met Ser Arg Leu Gly Arg Glu Asp Thr Ile Lys Ile Leu Arg Asp Ala Gly Tyr His Val Lys Ala Asn Gly Val 95

GAT GTA ACA ACA CAT CGT CAA GAC ATT AAT GGA AAA GAA ATG AAA TTT GAA GTG TTA ACA TTG GCA AGC TTA ACA 375
Asp Val Thr Thr His Arg Gln Asp Ile Asn Gly Lys Glu Met Lys Phe Glu Val Leu Thr Leu Ala Ser Leu Thr 120

ACT GAA ATT CAA ATC AAC ATT GAG ATA GAA TCT AGA AAA TCC TAC AAA AAA ATG CTA AAA GAA ATC GGA GAG GTA 450
Thr Glu Ile Gln Ile Asn Ile Glu Ile Glu Ser Arg Lys Ser Tyr Lys Lys Met Leu Lys Glu Met Ala Glu Val 145

GCT CCA GAA TAC AGG CAT GAC TCT CCT GAT TGT GGG ATG ATA ATA TTA TGT ATA GCA GCA TTA GTA ATA ACT AAA 525
Ala Pro Glu Tyr Arg His Asp Ser Pro Asp Cys Gly Met Ile Ile Leu Cys Ile Ala Ala Leu Val Ile Thr Lys 170

TTA GCA GCA GGG GAC AGA TCT GGT CTT ACA GCC GTG ATT AGG AGA GCT AAT AAT GTC CTA AAA AAT GAA ATG AAA 600
Leu Ala Ala Gly Asp Arg Ser Gly Leu Thr Ala Val Ile Arg Arg Ala Asn Asn Val Leu Lys Asn Glu Met Lys 195

CGT TAC AAA GGC TTA CTA CCC AAG GAC ATA GCC AAC AGC TTC TAT GAA GTG TTT GAA AAA CAT CCC CAC TTT ATA 675
Arg Tyr Lys Gly Leu Leu Pro Lys Asp Ile Ala Asn Ser Phe Tyr Glu Val Phe Glu Lys His Pro His Phe Ile 220
```

CHART 16 (Continued)

```
GAT GTT TTT GTT CAT TTT GGT ATA GCA CAA TCT TCT ACC AGA GGT GGC AGT AGA GTT GAA GGG ATT TTT GCA GGA 750
Asp Val Phe Val His Phe Gly Ile Ala Gln Ser Ser Thr Arg Gly Gly Ser Arg Val Glu Gly Ile Phe Ala Gly 245

TTG TTT ATG AAT GCC TAT GGT GCA GGG CAA GTG ATG TTA CGG TGG GGA GTC TTA GCA AAA TCA GTT AAA AAT ATT 825
Leu Phe Met Asn Ala Tyr Gly Ala Gly Gln Val Met Leu Arg Trp Gly Val Leu Ala Lys Ser Val Lys Asn Ile 270

ATG TTA GGA CAT GCT AGT GTG CAA GCA GAA ATG GAA CAA GTT GTT GAG GTT TAT GAA TAT GCC CAA AAA TTG GGT 900
Met Leu Gly His Ala Ser Val Gln Ala Glu Met Glu Gln Val Val Glu Val Tyr Glu Tyr Ala Gln Lys Leu Gly 295

GGT GAA GCA GGA TTC TAC CAT ATA TTG AAC AAC CCA AAA GCA TCA TTA TTA TCT TTG ACT CAA TTT CCT CAC TTC 975
Gly Glu Ala Gly Phe Tyr His Ile Leu Asn Asn Pro Lys Ala Ser Leu Leu Ser Leu Thr Gln Phe Pro His Phe 320

TCC AGT GTA GTA TTA GGC AAT GCT GCT GGC CTA GGC ATA ATG GGA GAG TAC AGA GGT ACA CCG AGG AAT CAA GAT 1050
Ser Ser Val Val Leu Gly Asn Ala Ala Gly Leu Gly Ile Met Gly Glu Tyr Arg Gly Thr Pro Arg Asn Gln Asp 345

CTA TAT GAT GCA GCA AAG GCA TAT GCT GAA CAA CTC AAA GAA AAT GGT GTG ATT AAC TAC AGT GTA CTA GAC TTG 1125
Leu Tyr Asp Ala Ala Lys Ala Tyr Ala Glu Gln Leu Lys Glu Asn Gly Val Ile Asn Tyr Ser Val Leu Asp Leu 370

ACA GCA GAA GAA CTA GAG GCT ATC AAA CAT CAG CTT AAT CCA AAA GAT AAT GAT GTA GAG CTT TGA GTT AAT      1197
Thr Ala Glu Glu Leu Glu Ala Ile Lys His Gln Leu Asn Pro Lys Asp Asn Asp Val Glu Leu                  301
```

CLAIMS

We claim:

1. A plasmid comprising a DNA sequence coding for human respiratory syncytial virus structural proteins selected from the group consisting of:

a. F protein;

b. G protein;

c. 22 K protein;

d. 9.5 K protein;

e. Major capsid protein N; and,

immunogenic fragments thereof.

2. The plasmid of claim 1 wherein the plasmid facilitates expression of the DNA sequence coding for human respiratory syncytial virus structural proteins by a suitable host.

3. The plasmid of claim 2 wherein the expression of the human respiratory syncytial virus structural genes are under the control of a cytomegalovirus promoter.

4. The plasmid of claim 2 wherein the replication of the plasmid while in a suitable eukaryote host is under the control of bovine papilloma virus DNA sequences.

5. A recombinant virus of the baculovirus family capable of expressing a DNA sequence coding for human respiratory syncytial virus structural proteins selected from the group consisting of:

a. F protein;

b. G protein;

c. 22 K protein;

d. 9.5 K protein;

e. Major capsid protein N; and,

immunogenic fragments thereof.

6. The virus of claim 5 wherein the virus is Autographa californica nuclear polyhedral virus.

7. A plasmid according to claim 2 wherein the suitable host is selected from among the group comprising:

a. bacteria;

b. yeast; and,

c. eukaryote cell cultures.

8. A suitable host containing a plasmid comprising a DNA sequence coding for human respiratory syncytial virus structural proteins selected from the group consisting of:

a. F protein;

b. G protein;

c. 22 K protein;

d. 9.5 K protein;

e. Major capsid protein N; and,

immunogenic fragments thereof.

9. A suitable host according to claim 8 wherein the plasmid facilitates expression of the DNA sequence coding for human respiratory syncytial virus structural proteins.

10. A suitable host according to claim 8 wherein the host is selected from the group comprising:

a. bacteria;

b. yeast; and,

c. eukaryote cell cultures.

11. A suitable host according to claim 9 wherein the host is selected from the group comprising:

a. bacteria;

b. yeast; and,

c. eukaryote cell cultures.

12. An essentially pure protein selected from the group consisting of:

a. F protein;

b. G protein;

c. 22 K protein;

d. 9.5 K protein;

e. Major capsid protein N; and,
immunogenic fragments thereof.

13. A human respiratory syncytial virus vaccine comprising one or more essentially pure proteins selected from the group consisting of:

a. F protein;
b. G protein;
c. 22 K protein;
d. 9.5 K protein;
e. Major capsid protein N; and,
immunogenic fragments thereof.

14. A method for protecting humans from human respiratory syncytial virus by vaccination with a vaccine comprising one or more essentially pure proteins selected from the group consisting of:

a. F protein;
b. G protein;
c. 22 K protein;
d. 9.5 K protein;
e. Major capsid protein N; and,
immunogenic fragments thereof.

INTERNATIONAL SEARCH REPORT

International Application No PCT/US 86/02756

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [6]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC[4]: C 12 N 15/00; C 12 N 7/00; C 12 N 1/20; C 12 N 1/18; C 12 N 5/00; C 07 K 15/04; A 61 K 39/155

II. FIELDS SEARCHED

Minimum Documentation Searched [7]

Classification System	Classification Symbols
IPC[4]	C 12 N; A 61 K

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

Category *	Citation of Document, [11] with indication, where appropriate, of the relevant passages [12]	Relevant to Claim No. [13]
X	Proceedings of the National Academy of Sciences of the USA, volume 81, December 1984, P.L. Collins et al.: "Nucleotide sequence of the gene encoding the fusion (F) glycoprotein of human respiratory syncytial virus", pages 7683-7687 see the whole document cited in the application	1,8
Y		5,6
	--	
Y	Chemical Abstracts, volume 104, no. 21, 26 May 1986, (Columbus, Ohio, US), M.D. Summers et al.: "Genetic engineering of the genome of the Autographa californica nuclear polyhedrosis virus", see page 125, abstract 180754w, & Banbury Rep. 1985, 22(Genet. Altered Viruses Environ.), 319-39	5,6
	--	
	./.	

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
26th May 1987	19 JUN 1987
International Searching Authority	**Signature of Authorized Officer**
EUROPEAN PATENT OFFICE	M. VAN MOL

FURTHER INFORMATION CONTINUED FROM THE SECOND SHEET

X	Proceedings of the National Academy of Sciences of the USA, volume 82, June 1985, G.W. Wertz et al.: "Nucleotide sequence of the G protein gene of human respiratory syncytial virus reveals an unusual type of viral membrane protein", pages 4075-4079 see the whole document cited in the application -- ./..	1,8

V. [X] OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE [1]

This International search report has not been established in respect of certain claims under Article 17(2) (a) for the following reasons:

1. [X] Claim numbers 14, because they relate to subject matter not required to be searched by this Authority, namely:

See PCT Rule 39.1(iv):
Methods for treatment of the human or animal body by means of surgery or therapy, as well as diagnostic methods.

2. [ ] Claim numbers ....., because they relate to parts of the International application that do not comply with the prescribed requirements to such an extent that no meaningful International search can be carried out, specifically:

3. [ ] Claim numbers ....., because they are dependent claims and are not drafted in accordance with the second and third sentences of PCT Rule 6.4(a).

VI. [ ] OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING [2]

This International Searching Authority found multiple inventions in this International application as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International search report covers all searchable claims of the International application.

2. [ ] As only some of the required additional search fees were timely paid by the applicant, this International search report covers only those claims of the International application for which fees were paid, specifically claims:

3. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:

4. [ ] As all searchable claims could be searched without effort justifying an additional fee, the International Searching Authority did not invite payment of any additional fee.

Remark on Protest

[ ] The additional search fees were accompanied by applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category *	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No
X	Journal of Virology, volume 54, no. 1, April 1985, American Society for Microbiology, P.L. Collins et al.: "The envelope-associated 22K protein of human respiratory syncytial virus: Nucleotide sequence of the mRNA and a related polytranscript", pages 65-71 see the abstract cited in the application	1,8
X	Virology, volume 146, no. 1, January 1985, Academic Press, Inc., P.L. Collins et al.: "Correct sequence for the major nucleocapsid protein mRNA of respiratory syncytial virus", pages 69-77 see the abstract cited in the application	1,8